

STIC-Biotech/ChemLib

169038

Mej

From: Swope, Sheridan
Sent: Wednesday, October 19, 2005 5:09 PM
To: STIC-Biotech/ChemLib
Subject: 10/698,228

For 10/698,228, pls do the following:

Align SID 1 with SID 2
Align SID 1 with SID 12

Search and Interference Search:

SID 1 against the NT and AA data bases
SID 2 against the NT and AA data bases
SID 12 against the NT and AA data bases

Thanks!!

Sheridan Swope, Ph.D.
Patent Examiner, AU 1656
Recombinant Enzymes
571-272-0943 (voice)
E02B71 Remsen Bld (Office)
E03C70 Remsen Bld (Mailbox)

Toby Port

RECEIVED
OCT 20 2005
STIC/CHEN, LIA
(STIC)

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

STIC-Biotech/ChemLib

166062

M2

From: Swope, Sheridan
Sent: Saturday, September 17, 2005 3:04 PM
To: STIC-Biotech/ChemLib
Subject: 10/698,228

For 10/698,228,

Pls search and interference search:
SID 12 against the NT data bases.

Pls, also align SID 2 and SID 12

Sheridan Swope, Ph.D.
Patent Examiner, AU 1656
Recombinant Enzymes
571-272-0943 (voice)
E02B71 Remsen Bld (Office)
E03C70 Remsen Bld (Mailbox)

12 na 1053

2 na 1053

Bret Sharts

M2

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

170611
Shears, Beverly

From: Swope, Sheridan
Sent: Thursday, November 03, 2005 4:17 PM
To: Shears, Beverly
Subject: FW: 10/698,228

Importance: High

Beverly,

The results I got back indicate that the search was done for sid 2, not sid 12.
Would you do the search for sid 12 asap?

thanks,

Sheridan

-----Original Message-----

From: Swope, Sheridan
Sent: Saturday, September 17, 2005 3:04 PM
To: STIC-Biotech/ChemLib
Subject: 10/698,228

Bey Shears

For 10/698,228,

Pls search and interference search:
SID 12 against the NT data bases.

Pls, also align SID 2 and SID 12

Sheridan Swope, Ph.D.
Patent Examiner, AU 1656
Recombinant Enzymes
571-272-0943 (voice)
E02B71 Remsen Bld (Office)
E03C70 Remsen Bld (Mailbox)

19280

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2005, 06:51:46 ; Search time 672 Seconds
(without alignments)
9276.023 Million cell updates/sec

Title: US-10-698-228-2

Perfect score: 1053

Sequence: 1 atggcgaccggaaagcc.....tcacottgagcagatttt 1053

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002as:*

7: Geneseqn2002bs:*

8: Geneseqn2003as:*

9: Geneseqn2003bs:*

10: Geneseqn2003cs:*

11: Geneseqn2003ds:*

12: Geneseqn2004as:*

13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1053	100.0	1053	4	Aaf32438 Human rib
2	1053	100.0	2596	3	Aa12411 cDNA enco
3	1053	100.0	4955	4	Aaf32439 Human rib
4	1053	100.0	4955	13	Adr24210 Breast ca
5	1051.4	99.8	1053	4	Aaf32447 Human rib
6	1051.4	99.8	1081	4	Aaf32440 Human rib
7	1049.8	99.7	1601	4	Aa114924 Human cDN
8	650.6	61.8	706	13	Adg57092 Novel can
9	578.4	54.9	1989	3	Aac78111 Human can
10	578.4	54.9	2215	12	Adk70302 Respirato
11	578.4	54.9	2216	10	Adj56536 Murine cD
12	578.4	54.9	2482	4	Aas44917 Human con
13	578.4	54.9	2500	6	Ab165414 Lung canc
14	578.4	54.9	2500	6	Ab166517 Lung canc
15	578.4	54.9	2500	6	Ab165859 Lung canc
16	578.4	54.9	2500	8	Abx10335 DNA enco
17	578.4	54.9	2500	11	Adi32044 Human cDN
18	578.4	54.9	2500	12	Adn04443 Antipsoi
19	578.4	54.9	2500	12	Adg09273 Human RRM
20	578.4	54.9	2500	13	Acn37637 Tumour-as

	21	578.4	54.9	3294	4	AAH73225	Aah73225 Human cer
	22	578.4	54.9	3393	12	ADN03788	Adn03788 Antipsoi
	23	578.4	54.9	3393	12	ADO19225	Ado19225 Human PRO
	24	569	54.0	2641	4	AAS44745	Aas44745 Human ful
	25	568.8	54.0	1170	2	AAV05641	Aav05641 Human rib
	26	561	53.3	1328	12	ADO57308	Ado57308 DNA enco
	27	547.8	52.0	978	4	AAS41006	Aas41006 cDNA enco
	28	545.4	51.8	977	6	ABL90228	Ab190228 Human pol
	29	538.8	50.2	608	4	AAH07707	Aah07707 Human cDN
	30	514	48.8	1371	5	AAS79474	Aas79474 DNA enco
	31	463	44.0	1289	4	ABL14627	Ab114627 Drosophil
	32	442	42.0	1146	6	AAS62387	Aas62387 cDNA sequ
	33	441	41.9	481	2	AAS51874	Aas51874 Human sec
C	34	432.2	41.0	1218	8	ABT20895	Abt20895 Aspergill
	35	429.8	40.8	14176	2	RAT84564	Rat84564 Swinepox
	36	429.8	40.8	14175	4	AAF84949	Aaf84949 Reverse c
C	37	429.8	40.8	14176	4	AAF84948	Aaf84948 Nucleotid
	38	420.4	39.9	3945	4	ABL14626	Ab114626 Drosophil
	39	403.2	38.3	963	13	ADS58378	Ads58378 Bacterial
	40	397	37.7	1146	8	ABT19075	Abt19075 Aspergill
	41	392.6	37.3	1173	13	ADT47814	Adt47814 Bacterial
	42	392	37.2	1242	6	ABZ32250	Candida a
	43	386.6	36.7	1292	13	ADS49379	Ads49379 Bacterial
	44	380.4	36.1	1200	10	ACC61238	Acc61238 Gene sequ
	45	380.4	36.1	1200	10	ADK63241	Adk63241 Disease t

ALIGNMENTS

RESULT 1

AAF32438

ID AAF32438 standard; cDNA; 1053 BP.

XX AAF32438;

AC AAF32438;

XX 18-APR-2001 (first entry)

DT Human ribonucleotide reductase TP53R2H nucleotide sequence SEQ ID NO:2.

DE Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.

XX Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.

XX Homo sapiens.

XX WO200100799-A1.

XX 04-JAN-2001.

XX 27-JUN-2000; 2000WO-JP004189.

XX 28-JUN-1999; 99JP-00181131.

PR 06-JUL-1999; 99JP-00192391.

PR 21-JAN-2000; 2000JP-00017770.

XX (TAKE) TAKEDA CHEM IND LTD.

XX (NAKA/) NAKAMURA Y.

PI Nakamura Y, Arakawa H, Tanaka H;

XX WPI; 2001-112446/12.

DR P-PSDB; AAB69050.

XX Ribonucleotide reductase involved in DNA repair and DNA encoding it, for

PT diagnosis, treatment and prevention of cancer.

XX Claim 5; Fig 1-3; 102pp; Japanese.

CC The present sequence encodes a human ribonucleotide reductase designated

CC TP53R2H. TP53R2H has cytosolic activity. TP53R2H is part of the DNA

CC repair mechanism and its activity is induced by p53. It can be used for

CC the treatment, prevention and diagnosis of a wide range of cancers

XX Sequence 1053 BP; 314 A; 192 C; 246 G; 301 T; 0 U; 0 Other;

		Query Match	100.0%; Score 1053; DB 4; Length 1053;	
		Best Local Similarity	100.0%; Pred. NO. 1e-290;	
		Mismatches 1053; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	ATGGGCGACCCGGAAGCGCGGAGCGGGCTGGATCAGATCAGATCATCTTCA	60	
DB	1	ATGGGCGACCCGGAAGCGCGGAGCGGGCTGGATCAGATCAGATCATCTTCA	60	
QY	61	GACACCAACGAAGTGAATAAAGTCAATGAAGCCACTCTCTAAGAAAGAGTTCTCGC	120	
DB	61	GACACCAACGAAGTGAATAAAGTCAATGAAGCCACTCTCTAAGAAAGAGTTCTCGC	120	
QY	121	CGGTTTGTGCATCTTTCCATCCAGTACCCTGATATTTGGAAATGATATAACAGGCACAG	180	
DB	121	CGGTTTGTGCATCTTTCCATCCAGTACCCTGATATTTGGAAATGATATAACAGGCACAG	180	
QY	181	GCTTCTCTTGACAGCAGAGGTCGACTATCAAGATCTCCCTCAGTGAACAAG	240	
DB	181	GCTTCTCTTGACAGCAGAGGTCGACTATCAAGATCTCCCTCAGTGAACAAG	240	
QY	241	CTTAAAGCAGATGAGAAGTACTCTCTCTCATCTTACATCTTACGCTTTTTCAGCCAGTGAT	300	
DB	241	CTTAAAGCAGATGAGAAGTACTCTCTCTCATCTTACATCTTACGCTTTTTCAGCCAGTGAT	300	
QY	301	GGAAATTGTAATCAAAATTTGGTGGAGCGCTTTTAGTCAGGAGGTGAGGTTCCAGAGGCT	360	
DB	301	GGAAATTGTAATCAAAATTTGGTGGAGCGCTTTTAGTCAGGAGGTGAGGTTCCAGAGGCT	360	
QY	361	CGCTGTTTCTATGGCTTCAAAATTCATCGAGAATGTTCACTCAGAGATGACAGTTTG	420	
DB	361	CGCTGTTTCTATGGCTTCAAAATTCATCGAGAATGTTCACTCAGAGATGACAGTTTG	420	
QY	421	CTGATAGACATTCATCAGAGATCCCAAGAAAGGAAATTTTAAATCAATTTGAA	480	
DB	421	CTGATAGACATTCATCAGAGATCCCAAGAAAGGAAATTTTAAATCAATTTGAA	480	
QY	481	ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGATGAGTAGCAGATAGAAAA	540	
DB	481	ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGATGAGTAGCAGATAGAAAA	540	
QY	541	TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAAGAGAGTTTCTTCTCAGGA	600	
DB	541	TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAAGAGAGTTTCTTCTCAGGA	600	
QY	601	TCTTTTGTCTATATCTGGCTTAAGAGAGAGGTTCTATGCGAGACTCACTTTTTC	660	
DB	601	TCTTTTGTCTATATCTGGCTTAAGAGAGAGGTTCTATGCGAGACTCACTTTTTC	660	
QY	661	AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTTGTGCTGATGTTCCAA	720	
DB	661	AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTTGTGCTGATGTTCCAA	720	
QY	721	TACTTAGTAATAAGCTTTCAGAAAGAGGTGAGGAGATCATTTGATGCTGCAAA	780	
DB	721	TACTTAGTAATAAGCTTTCAGAAAGAGGTGAGGAGATCATTTGATGCTGCAAA	780	
QY	781	ATTGACGAGAGTTTAAACAGAGCTTCCAGTTGGCTCATTTGGAATGAATTCATT	840	
DB	781	ATTGACGAGAGTTTAAACAGAGCTTCCAGTTGGCTCATTTGGAATGAATTCATT	840	
QY	841	TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGAACCTTGGATTCTCA	900	
DB	841	TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGAACCTTGGATTCTCA	900	
QY	901	AAGGTTTTTTCAGGCAAGAAATCTTTTGAATTTTATGAAACATTTCTTTAGAGGAAA	960	
DB	901	AAGGTTTTTTCAGGCAAGAAATCTTTTGAATTTTATGAAACATTTCTTTAGAGGAAA	960	
QY	961	ACAAATTTCTTTGAGAAACAGGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC	1020	
DB	961	ACAAATTTCTTTGAGAAACAGGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC	1020	

QY	1021	ACAGATAACGTCCTTCACTTGGATCGAGATTTT	1053
DB	1021	ACAGATAACGTCCTTCACTTGGATCGAGATTTT	1053
RESULT 2			
AA12411	ID	AAA12411 standard; cDNA; 2596 BP.	
XX	AC	AAA12411;	
XX	25-JUL-2000	(first entry)	
DE	cDNA encoding a human RNA-associated protein.		
XX	Human; RNA-associated protein; cell proliferation; immune response; reproductive disorder; actinic keratosis; arteriosclerosis; bursitis; cirrhosis; mixed connective tissue disease; myelofibrosis; paroxysmal nocturnal hemoglobinuria; polycythemia vera; psoriasis; trauma; ss.		
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	56..1111	
FT	/*tag= a		
FT	/product= "RNA-associated protein"		
XX			
PN	WO200015799-A2.		
XX			
PD	23-MAR-2000.		
XX			
PF	17-SEP-1999; 99WO-US021688.		
XX			
PR	17-SEP-1998; 98US-00156039.		
PR	22-SEP-1998; 98US-00158720.		
PR	04-NOV-1998; 98US-00186815.		
PR	08-APR-1999; 99US-0128660P.		
XX	(INCY-) INCYTE PHARM INC.		
XX			
PI	Tang YT, Corley NC, Guegler KJ, Gorgone GA, Patterson C;		
PI	Hillman JL, Baughn MR, Lal P, Azimzai Y, Yue H, Yang J;		
XX			
DR	WPI: 2000-271437/23.		
DR	P-PSDB; AAY84439.		
XX			
PT	New polypeptides and polynucleotides, useful for preventing and treating a disorder associated with increased or decreased expression of RNA		
PT	associated proteins.		
XX			
PS	Claim 9; Page 120-121; 131pp; English.		
XX			
CC	The present sequence encodes a human RNA-associated protein. The expression of RNA-associated proteins is closely associated with reproductive tissues, nervous tissues, cell proliferation including cancer, inflammation and immune responses, and so they may be used for diagnosis, treatment or prevention of cell proliferative,		
CC	immune/inflammatory disorders, and reproductive disorders. Diseases and disorders which may be treated include actinic keratosis,		
CC	atherosclerosis, arteriosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease, myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia		
CC	and cancers, and trauma		
SQ	Sequence 2596 BP; 783 A; 418 C; 510 G; 885 T; 0 U; 0 Other;		
Query Match 100.0%; Score 1053; DB 3; Length 2596;			
Best Local Similarity 100.0%; Pred. No. 1.5e-290;			
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1 ATGGGCGACCCGGAAGCGCGGAGCGGGCTGGATCAGATCAGATCATCTTCA 60		

```
Db 56 ATGGGCGACCCGGAAGGCGGGAAGCGCGGCTGGATCAGATGAGATCATCTTCA 115
Qy 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 116 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 175
Qy 121 CGGTTTGTCTATCTTCCATCCAGTACCCTCGATATTGGAAATGATATAACAGGCGACAG 180
Db 176 CGGTTTGTCTATCTTCCATCCAGTACCCTCGATATTGGAAATGATATAACAGGCGACAG 235
Qy 181 GCTTCCCTTCTGGACAGCAGAAAGAGTGCACCTTATCAAGAGATCTCCCTCACTGCAACAAG 240
Db 236 GCTTCCCTTCTGGACAGCAGAAAGAGTGCACCTTATCAAGAGATCTCCCTCACTGCAACAAG 295
Qy 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCACATCTTAGCCTTTTGGCAGCCAGTGAT 300
Db 296 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCACATCTTAGCCTTTTGGCAGCCAGTGAT 355
Qy 301 GGAATTTGTAATAAGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
Db 356 GGAATTTGTAATAAGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 415
Qy 361 CGCTGTTTCTATGGCTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db 416 CGCTGTTTCTATGGCTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 475
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAATTTTATTAATGCAATTGAA 480
Db 476 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAATTTTATTAATGCAATTGAA 535
Qy 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCCCTTGCATGATAGCAGATAGAAAA 540
Db 536 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCCCTTGCATGATAGCAGATAGAAAA 595
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAGAGATTTTCTTCTCAGGA 600
Db 596 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAGAGATTTTCTTCTCAGGA 655
Qy 601 TCTTTTGTCTATATCTGGCTAAAGAGAGAGTCTTATGCGAGACTCACTTTTTC 660
Db 656 TCTTTTGTCTATATCTGGCTAAAGAGAGAGTCTTATGCGAGACTCACTTTTTC 715
Qy 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 716 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 775
Qy 721 TACTTAGTAAATAAGCCTTCAGAAAGAAAGGTCAGGAGATCAATGTTGATGCTGTCAAA 780
Db 776 TACTTAGTAAATAAGCCTTCAGAAAGAAAGGTCAGGAGATCAATGTTGATGCTGTCAAA 835
Qy 781 ATTGAGCAGAGTTTTAAACAGAGAGCTTGCAGTGGCTCATTTGGAATGAATTCATTT 840
Db 836 ATTGAGCAGAGTTTTAAACAGAGAGCTTGCAGTGGCTCATTTGGAATGAATTCATTT 895
Qy 841 TTGATGAACACAGTACATGAGTTTGTAGCTGCAGAGATTAATCTGTGAACTTTGATTTCA 900
Db 896 TTGATGAACACAGTACATGAGTTTGTAGCTGCAGAGATTAATCTGTGAACTTTGATTTCA 955
Qy 901 AAGGTTTTTTCAGGCAGAAAATCCTTTTGTATTTTATGGAACAACTTTCTTTAGAGGAATA 960
Db 956 AAGGTTTTTTCAGGCAGAAAATCCTTTTGTATTTTATGGAACAACTTTCTTTAGAGGAATA 1015
Qy 961 ACAATTTCTTTGAGAAAACAGTTTTCAGAGTATCAGGTTTTCAGGTTTTCAGGAGAAACC 1020
Db 1016 ACAATTTCTTTGAGAAAACAGTTTTCAGAGTATCAGGTTTTCAGGTTTTCAGGAGAAACC 1075
Qy 1021 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1053
Db 1076 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1108
```

RESULT 3

```
AAF32439
ID AAF32439 standard; cDNA; 4955 BP.
XX
XX AAF32439;
AC
XX
DT 18-APR-2001 (first entry)
XX
DE Human ribonucleotide reductase related nucleotide sequence SEQ ID NO:3.
XX
XX Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.
XX
XX Homo sapiens.
XX
XX WO200100799-A1.
XX
XX 04-JAN-2001.
XX
XX 27-JUN-2000; 2000WO-JP004189.
XX
XX 28-JUN-1999; 99JP-00181131.
XX
XX 06-JUL-1999; 99JP-00192391.
XX
XX 21-JAN-2000; 2000JP-00017770.
XX
XX (TAKE ) TAKEDA CHEM IND LTD.
XX
XX (NAKA/) NAKAMURA Y.
XX
XX Nakamura Y, Arakawa H, Tanaka H;
XX
XX WPI; 2001-112446/12.
XX
XX Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
XX
XX diagnosis, treatment and prevention of cancer.
XX
XX Example 2; Page 87-90; 102pp; Japanese.
XX
XX The present invention describes a human ribonucleotide reductase
XX
XX designated TP53R2H. TP53R2H has cytosolic activity. TP53R2H is part of
XX
XX the DNA repair mechanism and its activity is induced by p53. It can be
XX
XX used for the treatment, prevention and diagnosis of a wide range of
XX
XX cancers. The present sequence represents a human ribonucleotide reductase
XX
XX related sequence which is used in an example from the present invention
XX
XX
XX Sequence 4955 BP; 1526 A; 803 C; 999 G; 1627 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 1053; DB 4; Length 4955;
XX
XX Best Local Similarity 100.0%; Pred. No. 2.1e-290;
XX
XX Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
Qy 1 ATGGGCGACCCGGAAGGCGGGAAGCGCGGCTGGATCAGATGAGATCATCTTCA 60
Db 245 ATGGGCGACCCGGAAGGCGGGAAGCGCGGCTGGATCAGATGAGATCATCTTCA 304
Qy 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 305 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 364
Qy 121 CGGTTTGTCTATCTTCCATCCAGTACCCTCGATATTGGAAATGATATAACAGGCGACAG 180
Db 365 CGGTTTGTCTATCTTCCATCCAGTACCCTCGATATTGGAAATGATATAACAGGCGACAG 424
Qy 181 GCTTCCCTTCTGGACAGCAGAAAGAGTGCACCTTATCAAGAGATCTCCCTCACTGCAACAAG 240
Db 425 GCTTCCCTTCTGGACAGCAGAAAGAGTGCACCTTATCAAGAGATCTCCCTCACTGCAACAAG 484
Qy 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCACATCTTTAGCCTTTTGGCAGCCAGTGAT 300
Db 485 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCACATCTTTAGCCTTTTGGCAGCCAGTGAT 544
Qy 301 GGAATTTGTAATAAGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTTCCAGAGGCT 360
Db 545 GGAATTTGTAATAAGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTTCCAGAGGCT 604
Qy 361 CGCTGTTTCTATGGCTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
```

Db 605 CCGTGTTCATATGCGCTTCAAAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTG 664
 Qy 421 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTTGAA 480
 Db 665 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTTGAA 724
 Qy 481 ACATGCCCTATGTTAAGAAAAAAGCAGATTTGGCCCTTCGATGGATAGCAGATAGAAAA 540
 Db 725 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGCCCTTCGATGGATAGCAGATAGAAAA 784
 Qy 541 TCTACTTTTGGGAAAGAGTGGGCTTTGCTGCTAGAGAGTGTCTTCTCAGGA 600
 Db 785 TCTACTTTTGGGAAAGAGTGGGCTTTGCTGCTAGAGAGTGTCTTCTCAGGA 844
 Qy 601 TCTTTTGTCTATATTTCTGGCTTAAAGAAAGAGAGTGTCTTATGCGAGGACACTTTTTC 660
 Db 845 TCTTTTGTCTATATTTCTGGCTTAAAGAAAGAGAGTGTCTTATGCGAGGACACTTTTTC 904
 Qy 661 AATGAATCATCAGCAGAGATGAAGACTTCACTGTGACTTTTGCCTGTGATGTTCCAA 720
 Db 905 AATGAATCATCAGCAGAGATGAAGACTTCACTGTGACTTTTGCCTGTGATGTTCCAA 964
 Qy 721 TACTTAGTAATAAGCTTTCAGAAAGAGTGTGCTGCTAGAGAGTGTCTTCTCAGGA 780
 Db 965 TACTTAGTAATAAGCTTTCAGAAAGAGTGTGCTGCTAGAGAGTGTCTTCTCAGGA 1024
 Qy 781 ATTGACGAGAGTGTAAACAGAGAGCTTCCAGTTGGCTCAATTTGGAATGAATTTGCATT 840
 Db 1025 ATTGACGAGAGTGTAAACAGAGAGCTTCCAGTTGGCTCAATTTGGAATGAATTTGCATT 1084
 Qy 841 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGAGTACTTTGGAATTTGGAATTTCTCA 900
 Db 1085 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGAGTACTTTGGAATTTGGAATTTCTCA 1144
 Qy 901 AAGGTTTTTTCAGCAGAGAAATCCTTTTGAATTTTATGGAACCAATTTCTTTAGAGGAAAA 960
 Db 1145 AAGGTTTTTTCAGCAGAGAAATCCTTTTGAATTTTATGGAACCAATTTCTTTAGAGGAAAA 1204
 Qy 961 ACAATTTTCTTTAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGCTATGCGAGAAACC 1020
 Db 1205 ACAATTTTCTTTAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGCTATGCGAGAAACC 1264
 Qy 1021 ACAGATAACGCTTTCACCTTTGGATGCGAGATTTT 1053
 Db 1265 ACAGATAACGCTTTCACCTTTGGATGCGAGATTTT 1297

RESULT 4
 ADR24210
 ID ADR24210 standard; DNA; 4955 BP.
 XX AC ADR24210;
 XX DT 21-OCT-2004 (first entry)
 XX DE Breast cancer prognosis marker #71.
 XX KW ds; breast cancer; prognosis; gene expression; diagnosis.
 XX OS Homo sapiens.
 XX PN WO2004065545-A2.
 XX PD 05-AUG-2004.
 XX PF 15-JAN-2004; 2004WO-US0001100.
 XX PR 15-JAN-2003; 2003US-00342887.
 XX PA (ROSE-) ROSETTA INPHARMATICS LLC.
 XX PA (NECA-) NETHERLANDS CANCER INST.

PI Van't Veer LJ, He Y;
 DR WPI; 2004-593473/57.
 XX Classifying a breast cancer patient according to prognosis comprises
 PT determining the similarity between the level of expression of each of
 PT five genes in a cell sample taken from patient, to control levels.
 XX Disclosure; SEQ ID NO 71; 226pp; English.
 XX The invention relates to a method of classifying a breast cancer patient
 CC according to prognosis by determining the similarity between the level of
 CC expression of each of five genes for which markers are listed in the
 CC specification, in a cell sample taken from the breast cancer patient, to
 CC control levels of expression for each respective five genes to obtain a
 CC patient similarity value. The methods are useful for classifying a breast
 CC cancer patient according to prognosis. Kits and computer program products
 CC are useful for data analysis using the diagnostic, prognostic and
 CC statistical methods of the invention. This sequence corresponds to a
 CC marker used in the method of the invention.
 XX SQ Sequence 4955 BP; 1526 A; 803 C; 999 G; 1627 T; 0 U; 0 Other;
 Query Match 100.0%; Score 1053; DB 13; Length 4955;
 Best Local Similarity 100.0%; Pred. No. 2.1e-290;
 Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 ATGGCGGACCGGAAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
 Db 245 ATGGCGGACCGGAAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
 Qy 61 GACACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCG 120
 Db 305 GACACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCG 364
 Qy 121 CGGTTGTCTATCTTCCAAATCCAGTACCTGATTTTGGAAAATGTATAAACAGGCACAG 180
 Db 365 CGGTTGTCTATCTTCCAAATCCAGTACCTGATTTTGGAAAATGTATAAACAGGCACAG 424
 Qy 181 GCTTCTCTTCGACAGAGAGAGTGCATTTATCAAAAGGATCTCCCTCAGTGAACAAG 240
 Db 425 GCTTCTCTTCGACAGAGAGAGTGCATTTATCAAAAGGATCTCCCTCAGTGAACAAG 484
 Qy 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCAGATCTTTAGCCTTTTTCAGCCAGTGTAT 300
 Db 485 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCAGATCTTTAGCCTTTTTCAGCCAGTGTAT 544
 Qy 301 GGAATTTGAATGAAAAATTTGGTGGAGCGCTTTTAGTCAGGAGTGCAGGTTCCAGAGGCT 360
 Db 545 GGAATTTGAATGAAAAATTTGGTGGAGCGCTTTTAGTCAGGAGTGCAGGTTCCAGAGGCT 604
 Qy 361 CGCTGTTTCTATGCTTCAAAATTTCTCATCGAGATGTTCACTCAGAGATGTACAGTTTG 420
 Db 605 CGCTGTTTCTATGCTTCAAAATTTCTCATCGAGATGTTCACTCAGAGATGTACAGTTTG 664
 Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTTGAA 480
 Db 565 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTTGAA 724
 Qy 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGCCCTTCGATGGATAGCAGATAGAAAA 540
 Db 725 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGCCCTTCGATGGATAGCAGATAGAAAA 784
 Qy 541 TCTACTTTTGGGAAAGAGTGGGCTTTGCTGCTGTAAGAGAGTGTCTTCTCAGGA 600
 Db 785 TCTACTTTTGGGAAAGAGTGGGCTTTGCTGCTGTAAGAGAGTGTCTTCTCAGGA 844
 Qy 601 TCTTTTGTCTATATTTCTGCTAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
 Db 845 TCTTTTGTCTATATTTCTGCTAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTTC 904
 Qy 661 AATGAATCATCAGCAGAGATGAAGACTTCACTGTGACTTTTGCCTGTGATGTTCCAA 720

Db 905 AATGAACATCATCAGCAGATGAAGGACTTCACCTGTGACTTTGCTTGCCTGATGTTCCAA 964
 Qy 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGTCAGGAGATCAATGTTGATGCTGTCAAA 780
 Db 965 TACTTAGTAAATAAGCCTTCAGAAAGAGGTCAGGAGATCAATGTTGATGCTGTCAAA 1024
 Qy 781 ATTGACAGAGGTTTTTAACAGAGAGCCTTCCAGTGTGGCCTCATTTGGAATGAATTGCATT 840
 Db 1025 ATTGACAGAGGTTTTTAACAGAGAGCCTTCCAGTGTGGCCTCATTTGGAATGAATTGCATT 1084
 Qy 841 TTGATGAACAGTACATGAGTTGTAGCTGACAGATTAATCTGTGAACTTGTGATTTCTCA 900
 Db 1085 TTGATGAACAGTACATGAGTTGTAGCTGACAGATTAATCTGTGAACTTGTGATTTCTCA 1144
 Qy 901 AAGGTTTTTTCAGGAGAGAAATCCTTTTGTGATTTTATGAGAAACATTTCTTTAGAGGAA 960
 Db 1145 AAGGTTTTTTCAGGAGAGAAATCCTTTTGTGATTTTATGAGAAACATTTCTTTAGAGGAA 1204
 Qy 961 ACAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGGTTATGCGAGAAACC 1020
 Db 1205 ACAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGGTTATGCGAGAAACC 1264
 Qy 1021 ACAGATAACGCTCTTACCTTGGATGCGAGATTTT 1053
 Db 1265 ACAGATAACGCTCTTACCTTGGATGCGAGATTTT 1297

RESULT 5

AAF32447
 ID AAF32447 standard; cDNA; 1053 BP.
 AC AAF32447;
 XX
 DT 18-APR-2001 (first entry)
 XX
 DE Human ribonucleotide reductase related nucleotide sequence SEQ ID NO:12.
 XX
 KW Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200100799-A1.
 XX
 PD 04-JAN-2001.
 XX
 PF 27-JUN-2000; 2000WO-JP004189.
 XX
 PR 28-JUN-1999; 99JP-00181131.
 PR 06-JUL-1999; 99JP-00192391.
 PR 21-JAN-2000; 2000JP-00017770.
 XX
 XX (TAKE) TAKEDA CHEM IND LTD.
 PA (NAKA/) NAKAMURA Y.
 PA
 XX Nakamura Y, Arakawa H, Tanaka H;
 PI WPI; 2001-112446/12.
 XX
 DR Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
 XX diagnosis, treatment and prevention of cancer.
 PT
 PT Claim 6; Page 96-97; 102pp; Japanese.
 PS
 CC The present invention describes a human ribonucleotide reductase
 CC designated TP53R2H. TP53R2H has cytosolic activity. TP53R2H is part of
 CC the DNA repair mechanism and its activity is induced by p53. It can be
 CC used for the treatment, prevention and diagnosis of a wide range of
 CC cancers. The present sequence represents a human ribonucleotide reductase
 CC related sequence which is given in the exemplification of the present
 CC invention
 CC
 XX Sequence 1053 BP; 314 A; 191 C; 246 G; 302 T; 0 U; 0 Other;

Query Match 99.8%; Score 1051.4; DB 4; Length 1053;
 Best Local Similarity 99.9%; Pred. No. 2.9e-290;
 Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATCAGAGATCATCTTCA 60
 Db 1 ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATCAGAGATCATCTTCA 60
 Qy 61 GACACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCTCTAAGAAAGATTTCTCG 120
 Db 61 GACACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCTCTAAGAAAGATTTCTCG 120
 Qy 121 CGGTTTGTCACTTTCATTCAGTACCTGATATTTGAAATGTATATAACAGGACAG 180
 Db 121 CGGTTTGTCACTTTCATTCAGTACCTGATATTTGAAATGTATATAACAGGACAG 180
 Qy 181 GCTTCTCTTCGACAGAGAGAGGTGACATTCATAAGAGGATCTCCCTCAGTGGAAACAG 240
 Db 181 GCTTCTCTTCGACAGAGAGAGGTGACATTCATAAGAGGATCTCCCTCAGTGGAAACAG 240
 Qy 241 CTTAAAGCAGATGAGAGTACTTCTCATCTCTCACATCTTTAGCCTTTTTCAGCCAGTAT 300
 Db 241 CTTAAAGCAGATGAGAGTACTTCTCATCTCTCACATCTTTAGCCTTTTTCAGCCAGTAT 300
 Qy 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAAGAGGTGAGGTTCCAGAGCT 360
 Db 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAAGAGGTGAGGTTCCAGAGCT 360
 Qy 361 CGCTGTTTCTATGCTTTCAAAATTTCTCATCAGAGATGTTCACTCAGAGATGTTACAGTTT 420
 Db 361 CGCTGTTTCTATGCTTTCAAAATTTCTCATCAGAGATGTTCACTCAGAGATGTTACAGTTT 420
 Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATCAATTTGAA 480
 Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATCAATTTGAA 480
 Qy 481 ACCATGCCCTATGTTAAGAAAGGAGATTTGGGCTTTGCGATGATGATAGCAGATAGAAA 540
 Db 481 ACCATGCCCTATGTTAAGAAAGGAGATTTGGGCTTTGCGATGATGATAGCAGATAGAAA 540
 Qy 541 TCTACTTTTGGGAAAGAGTGTGGCTTCTGCTGTAGAGAGGATTTTCTTCTCAGGA 600
 Db 541 TCTACTTTTGGGAAAGAGTGTGGCTTCTGCTGTAGAGAGGATTTTCTTCTCAGGA 600
 Qy 601 TCTTTTGTCTGCTATATTTCTGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTCC 660
 Db 601 TCTTTTGTCTGCTATATTTCTGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTCC 660
 Qy 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGTGCTGATTTCCAA 720
 Db 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGTGCTGATTTCCAA 720
 Qy 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGTCAGGAGATCAATTTGATGCTGTCAAA 780
 Db 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGTCAGGAGATCAATTTGATGCTGTCAAA 780
 Qy 781 ATTGACAGAGGTTTTTAACAGAGAGCCTTCCAGTGTGGCCTCATTTGGAATGAATTGCATT 840
 Db 781 ATTGACAGAGGTTTTTAACAGAGAGCCTTCCAGTGTGGCCTCATTTGGAATGAATTGCATT 840
 Qy 841 TTGATGAACAGTACATGAGTTGTAGCTGACAGATTAATCTGTGAACTTGTGATTTCTCA 900
 Db 841 TTGATGAACAGTACATGAGTTGTAGCTGACAGATTAATCTGTGAACTTGTGATTTCTCA 900
 Qy 901 AAGGTTTTTTCAGGAGAGAAATCCTTTTGTGATTTTATGAGAAACATTTCTTTAGAGGAAA 960
 Db 901 AAGGTTTTTTCAGGAGAGAAATCCTTTTGTGATTTTATGAGAAACATTTCTTTAGAGGAAA 960
 Qy 961 ACAATTTCTTTGAGAAACAGGTTTTCAGAGTATCAGCGTTTTCAGGTTATGCGAGAAACC 1020
 Db 961 ACAATTTCTTTGAGAAACAGGTTTTCAGAGTATCAGCGTTTTCAGGTTATGCGAGAAACC 1020
 Qy 1021 ACAGATAACGCTCTTACCTTGGATGCGAGATTTT 1053

Db	1021	ACAGATAACGCTTTCACCTTGGATCAGATTTT	1053
RESULT 6			
AAAF32440			
ID	AAAF32440	standard; cDNA; 1081 BP.	
XX	AAAF32440;		
DT	18-APR-2001	(first entry)	
XX	Human ribonucleotide reductase related nucleotide sequence SEQ ID NO:4.		
XX	Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.		
OS	Homo sapiens.		
XX	WO200100799-A1.		
PN	04-JAN-2001.		
PD	27-JUN-2000;	2000WO-JP004189.	
PF	28-JUN-1999;	99JP-00181131.	
PR	06-JUL-1999;	99JP-00192391.	
PR	21-JAN-2000;	2000JP-00017770.	
XX	(TAKE) TAKEDA CHEM IND LTD.		
PA	(NAKA/) NAKAMURA Y.		
XX	Nakamura Y, Arakawa H, Tanaka H;		
PI	WPI; 2001-112446/12.		
XX	Ribonucleotide reductase involved in DNA repair and DNA encoding it, for		
PT	diagnosis, treatment and prevention of cancer.		
XX	Example 2; Page 91; 102pp; Japanese.		
XX	The present invention describes a human ribonucleotide reductase		
CC	designated TP53R2H. TP53R2H has cytosolic activity. TP53R2H is part of		
CC	the DNA repair mechanism and its activity is induced by p53. It can be		
CC	used for the treatment, prevention and diagnosis of a wide range of		
CC	cancers. The present sequence represents a human ribonucleotide reductase		
CC	related sequence which is used in an example from the present invention		
XX	Sequence 1081 BP; 322 A; 199 C; 252 G; 308 T; 0 U; 0 Other;		
Query Match	99.8%;	Score 1051.4; DB 4; Length 1081;	
Best Local Similarity	99.9%;	Pred. No. 2.9e-290;	
Matches 1052;	Conservative 0;	Mismatches 1; Indels 0; Gaps 0;	
QY	1	ATGGGCGACCCGGAAGCGCGGCTGGATCAGATCAGATCATCTTCA	60
Db	20	ATGGGCGACCCGGAAGCGCGGCTGGATCAGATCAGATCATCTTCA	79
QY	61	GACACCAAGAAAGTGAATTAAGTGAAGAGCCACTCTTAAGAAAGTCTTCGC	120
Db	80	GACACCAAGAAAGTGAATTAAGTGAAGAGCCACTCTTAAGAAAGTCTTCGC	139
QY	121	CGGTTTGTTCATCTTCCATCCAGTACCTGATATTGGAAATGTATTAACAGGCACAG	180
Db	140	CGGTTTGTTCATCTTCCATCCAGTACCTGATATTGGAAATGTATTAACAGGCACAG	199
QY	181	GCTTCTCTTCGACAGCAGAGAGTCTGACTTATCAAGGATCTCCCTCACTGGAACAAG	240
Db	200	GCTTCTCTTCGACAGCAGAGAGTCTGACTTATCAAGGATCTCCCTCACTGGAACAAG	259
QY	241	CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCATCTCTTATGCGCCAGTGTAT	300
Db	260	CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCTCATCTCTTATGCGCCAGTGTAT	319
QY	301	GGAAATGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGCT	360
Db	320	GGAAATGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGCT	379
QY	361	CGCTGTTTCTATGGCTTTCAAAATTTCTCATCAGAAATGTTCACTCAGAGATGTACAGTTTG	420
Db	380	CGCTGTTTCTATGGCTTTCAAAATTTCTCATCAGAAATGTTCACTCAGAGATGTACAGTTTG	439
QY	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATGAA	480
Db	440	CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATGAA	499
QY	481	ACATGCCCTATGTTAAGAAAAGAGATGGCGCTTCCGATGGATGAGATAGAAAA	540
Db	500	ACATGCCCTATGTTAAGAAAAGAGATGGCGCTTCCGATGGATGAGATAGAAAA	559
QY	541	TCTACTTTTGGGAAAGAGTGGTGGCTTCTGCTGTAGAAAGAGTCTTCTCTCAGGA	600
Db	560	TCTACTTTTGGGAAAGAGTGGTGGCTTCTGCTGTAGAAAGAGTCTTCTCTCAGGA	619
QY	601	TCTTTTGTCTATATTCTGGCTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTCC	660
Db	620	TCTTTTGTCTATATTCTGGCTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTCC	679
QY	661	AATGAATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGGCTGATGTTCCAA	720
Db	680	AATGAATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGGCTGATGTTCCAA	739
QY	721	TACTTAGTAAATAGCCCTTCAGAAAGAAAGGCTCAGGAGATCATTTGATGCTGTCAA	780
Db	740	TACTTAGTAAATAGCCCTTCAGAAAGAAAGGCTCAGGAGATCATTTGATGCTGTCAA	799
QY	781	ATTGAGCAGGAGTCTTTTAAACAGAAAGCCCTTGGCAGTGGCCCTCATTTGGAATGAATTCATT	840
Db	800	ATTGAGCAGGAGTCTTTTAAACAGAAAGCCCTTGGCAGTGGCCCTCATTTGGAATGAATTCATT	859
QY	841	TTGATGAAAACAGTACATGAGTTTGTAGCTGACAGATTACTTTGGAACTTGGATCTCA	900
Db	860	TTGATGAAAACAGTACATGAGTTTGTAGCTGACAGATTACTTTGGAACTTGGATCTCA	919
QY	901	AAGGTTTTTTCAGGACAGAAATCCCTTTGATTTTATGGAACATTTCTTTAGAAGAAA	960
Db	920	AAGGTTTTTTCAGGACAGAAATCCCTTTGATTTTATGGAACATTTCTTTAGAAGAAA	979
QY	961	ACAAATTTCTTTGAGAAACAGAGTTTTCAGAGTATCAGCGTTTTCAGAGTTATGCGAGAAAC	1020
Db	980	ACAAATTTCTTTGAGAAACAGAGTTTTCAGAGTATCAGCGTTTTCAGAGTTATGCGAGAAAC	1039
QY	1021	ACAGATAACGCTTTCACCTTGGATGAGATTTT	1053
Db	1040	ACAGATAACGCTTTCACCTTGGATGAGATTTT	1072
RESULT 7			
AAH14924			
ID	AAH14924	standard; cDNA; 1601 BP.	
XX	AAH14924;		
XX	26-JUN-2001	(first entry)	
DT	Human cDNA sequence SEQ ID NO:12810.		
XX	Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.		
XX	Homo sapiens.		
OS	EP1074617-A2.		
XX	07-FEB-2001.		
PD	28-JUL-2000;	2000EP-00116126.	
PF			
XX			

PR 29-JUL-1999; 99JP-00248036.
 PR 27-AUG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.
 XX (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WPI; 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.
 XX
 PS Claim 8; SEQ ID NO 12810; 2537pp + Sequence Listing; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention
 XX
 SQ Sequence 1601 BP; 484 A; 290 C; 348 G; 479 T; 0 U; 0 Other;

Query Match 99.7%; Score 1049.8; DB 4; Length 1601;
 Best Local Similarity 99.8%; Pred. No. 1e-289;
 Matches 1051; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGGGACCCGGAAGCGCGGGCTGGATCAGGATGAGATCATCTTCA 60
 DB |||||
 QY 40 ATGGGGACCCGGAAGCGCGGGCTGGATCAGGATGAGATCATCTTCA 99
 DB |||||
 QY 61 GACACCAACGAAAGTGAATTAAGTCAATTAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
 DB |||||
 QY 100 GACACCAACGAAAGTGAATTAAGTCAATTAAGAGCCACTCTTAAGAAAGAGTTCTCGC 159
 QY 121 CGTTTGTCTATCTTCCATTCAGTACCTCGATATTTGGAAATGTATTAACAGGCACAG 180
 DB |||||
 QY 160 CGTTTGTCTATCTTCCATTCAGTACCTCGATATTTGGAAATGTATTAACAGGCACAG 219
 QY 181 GCCTTCCTTCGGACAGCAGAGAGGTGCTATCAAGAGTCTCCCTCACTCGAACAG 240
 DB |||||
 QY 220 GCCTTCCTTCGGACAGCAGAGAGGTGCTATCAAGAGTCTCCCTCACTCGAACAG 279
 QY 241 CTTTAAAGCAGATGAGAAGTACTTCTCATCTCTCATCTTAGCTTTTTCGAGCCAGTGAT 300
 DB |||||
 QY 280 CTTTAAAGCAGATGAGAAGTACTTCTCATCTCTCATCTTAGCTTTTTCGAGCCAGTGAT 339
 QY 301 GGAATTTGAATGAAATTTGGTGGAGCGCTTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 360
 DB |||||
 QY 340 GGAATTTGAATGAAATTTGGTGGAGCGCTTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 399

QY 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAAGTGTTCCTCAGAGATGTACAGTTTG 420
 DB |||||
 QY 400 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAAGTGTTCCTCAGAGATGTACAGTTTG 459
 DB |||||
 QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATCAATGAA 480
 DB |||||
 QY 460 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATCAATGAA 519
 DB |||||
 QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATTGGGCTTTCGATGATGATGATGATGATGAA 540
 DB |||||
 QY 520 ACCATGCCCTATGTTAAGAAAAGCAGATTGGGCTTTCGATGATGATGATGATGAA 579
 DB |||||
 QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTCGCTGTAGAGAGGATTTTCTCTCAGCA 600
 DB |||||
 QY 580 TCTACTTTTGGGAAAGAGTGGTGGCTTTCGCTGTAGAGAGGATTTTCTCTCAGCA 639
 DB |||||
 QY 601 TCTTTTGTCTATATCTGCTAAAGAGAGAGGTCTTATGCCAGGACTCATCTTTTTC 660
 DB |||||
 QY 640 TCTTTTGTCTATATCTGCTAAAGAGAGAGGTCTTATGCCAGGACTCATCTTTTTC 699
 DB |||||
 QY 661 AATGAATCATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
 DB |||||
 QY 700 AATGAATCATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 759
 DB |||||
 QY 721 TACTTAGTAATAAGCCTTTCAGAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAA 780
 DB |||||
 QY 760 TACTTAGTAATAAGCCTTTCAGAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAA 819
 DB |||||
 QY 781 ATTGAGCAGGAGTTTTAAACAGAGCCCTTGCAGTTGGCCTCATTTGGAATGAATTCATT 840
 DB |||||
 QY 820 ATTGAGCAGGAGTTTTAAACAGAGCCCTTGCAGTTGGCCTCATTTGGAATGAATTCATT 879
 DB |||||
 QY 841 TTGATGAACAAGTACATTTGAGTTTGTAGCTGACAGATTACTTTGTGAACTTTGATTTCTCA 900
 DB |||||
 QY 880 TTGATGAACAAGTACATTTGAGTTTGTAGCTGACAGATTACTTTGTGAACTTTGATTTCTCA 939
 DB |||||
 QY 901 AAGTTTTTTCAGGCAAGAAATCTTTTGAATTTTATGAAAACATTTCTTTAGAAGGAAA 960
 DB |||||
 QY 940 AAGTTTTTTCAGGCAAGAAATCTTTTGAATTTTATGAAAACATTTCTTTAGAAGGAAA 999
 DB |||||
 QY 961 ACAATTTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCGAGTTTATGCGAGAAACC 1020
 DB |||||
 QY 1000 ACAATTTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCGAGTTTATGCGAGAAACC 1059
 DB |||||
 QY 1021 ACAGATAACGCTTTCACCTTCGATGACAGATT 1053
 DB |||||
 QY 1060 ACAGATAACGCTTTCACCTTCGATGACAGATT 1092
 DB |||||

RESULT 8
 ADQ57092
 ID ADQ57092 standard; DNA; 706 BP.
 XX
 XX ADQ57092;
 XX AC AC
 XX 21-OCT-2004 (first entry)
 XX
 XX Novel canine microarray-related DNA sequence SeqID8394.
 DE canine microarray; drug screening; toxicity assay;
 XX environmental pollutant; cellular response; gene expression profile;
 KW toxic response; liver necrosis; fatty liver disease;
 KW protein adduct formation; hepatitis; dog; ds.
 XX
 OS Canis familiaris.
 XX
 XX WO2004063324-A2.
 XX
 XX 29-JUL-2004.
 PD
 XX 05-MAY-2003; 2003WO-US013853.
 XX
 XX 03-MAY-2002; 2002US-0377240P.
 PR

XX	(GENE-) GENE LOGIC INC.	
PA	(PFIZ) PFIZER PROD INC.	
XX		
PI	Diggins JC, Porter M, Wei T;	
XX	WPI; 2004-561890/54.	
DR		
XX	New isolated nucleic acid molecule, useful for drug screening and	
PT	toxicity assays or for assessing the impact, including toxicity, of a	
PT	compound, pharmaceutical agent or environmental pollutant on a cell or	
PT	living organism.	
XX		
PS	Claim 1; SEQ ID NO 8394; 41pp; English.	
XX		
CC	This invention is related to a novel isolated canine nucleic acid	
CC	sequences and the construction of canine microarrays containing a	
CC	significant portion of the canine genome. The isolated canine nucleic	
CC	acid sequences of the invention may be useful for drug screening and	
CC	toxicity assays. The invention is therefore useful for assessing the	
CC	impact, including toxicity, of a compound, pharmaceutical agent or	
CC	environmental pollutant on a cell or living organism. The methods are	
CC	useful for detecting genes that are up- or down-regulated in canines in a	
CC	disease state. The sequences are useful as diagnostic agents or markers	
CC	to detect a cellular response in a sample individually or as part of a	
CC	gene expression profile. It is also useful as a target for agents that	
CC	modulate gene expression or activity. The database is useful for	
CC	producing electronic Northernblots that allow the user to determine the cell	
CC	type or tissue in which a given gene is expressed and to allow	
CC	determination of the abundance or expression level of a given gene in a	
CC	particular tissue or cell. The methods are useful for determining the	
CC	similarity of a toxic response to one or more individual compounds. The	
CC	methods are useful for predicting at least one toxic response or the	
CC	likelihood that a compound or test agent will induce various specific	
CC	pathologies such as those of the liver (liver necrosis, fatty liver	
CC	disease, protein adduct formation or hepatitis), those of the kidney,	
CC	heart, brain or testes, or other pathologies associated with at least one	
CC	of the toxins. The methods are also useful for predicting or elucidating	
CC	the potential cellular pathways influenced, induced or modulated by the	
CC	compound or test agent due to the similarity of the expression profile	
CC	compared to the profile induced by a known toxin. The present sequence is	
CC	that of a canine DNA sequence which was claimed for use during the	
XX	production of a canine microarray of the invention.	
XX		
SQ	Sequence 706 BP; 208 A; 113 C; 158 G; 227 T; 0 U; 0 Other;	
	Query Match 61.8%; Score 650.6; DB 13; Length 706;	
	Best Local Similarity 95.2%; Pred. No. 1.4e-175;	
	Matches 671; Conservative 0; Mismatches 34; Indels 0; Gaps 0;	
QY	349 GTTCCAGAGGCTCGCTGTTTCTATGCTTTTCAATTTCTATCGAGAAATGTTCACTCAGAG	408
DB	1 GTTCCAGAGGCTCGCTGTTTCTATGCTTTTCAATTTCTATCGAGAAATGTTCACTCAGAG	60
QY	409 ATGTACAGTTTGTCTAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTT	468
DB	61 ATGTACAGTTTGTCTATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTT	120
QY	469 AATGCAATTGAAACCAATGCCCTATGTTAAGAAAAGGAGATTTGGCCCTTGCATGGATA	528
DB	121 AATGCAATTGAAACCAATGCCCTATGTTAAGAAAAGGAGATTTGGCCCTTGCATGGATA	180
QY	529 GCAGATAGAAAATCTACTTTTGGGGAAGAGTGGTGGCTTTCTCTGTAGAGAGATT	588
DB	181 GAAGATAGAAAATCTACTTTTGGGGAAGAGTGGTGGCTTTCTCTGTAGAGAGATT	240
QY	589 TTCTTCTCAGGATCTTTTGTCTGTATATCTTGCTTAAAGAGAGAGTCTTTATGCCAGGA	648
DB	241 TTCTTCTCAGGATCTTTTGTCTGTATATCTTGCTTAAAGAGAGAGTCTTTATGCCAGGA	300
QY	649 CTCACATTTTCCAAATGAATCTATCAGCAGAGATGAAGGAGTTCCACTGTGATTTGCTTGC	708
DB	301 CTCACATTTTCCAAATGAATCTATCAGCAGAGATGAAGGAGTTCCACTGTGATTTGCTTGC	360

709 CTGATGTTCCAAATCTAGTAAATAGCCCTTCAGAGAAAGGGTCAGGAGATCATTTCT 768

361 CTGATGTTCCAAATCTAGTAAATAGCCCTTCAGAGAAAGGGTCAGGAGATCATTTCT 420

769 GATGCTGTCAAATTCAGCAGAGGAGTTTAAACAGAACCTTGCCAGTTGGCCTCATTTGGA 828

421 AATGCTGTCAAATTCAGCAGAGGAGTTTAAACAGAACCTTGCCAGTTGGCCTCATTTGGA 480

829 ATGAATTTGCAATTTTGTATGAACACAGTACATTTGATGTTTGTAGCTGACAGATTAATTTGCGAA 888

481 ATGAATTTGCTTTTGTATGAACACAGTATATTTAGTTTGTAGCTGACAGATTAATTTGCGAA 540

889 CTTGGATTTCTCAAAGGTTTTCAGGCGAGAAATTCCTTTTGAATTTTATGGAACCAATTTCT 948

541 CTTGGATTTCTCAAAGGTTTTCAGGCGAGAAATTCCTTTTGAATTTTATGGAACCAATTTCT 600

949 TTAGAGGAAACCAATTTCTTTTGAACACAGTTCAGAGTATCAGCGTTTGTGCGATT 1008

601 TTAGAGGAAACCAATTTCTTTTGAACACAGTTCAGAGTATCAGCGTTTGTGCGATT 660

1009 ATGCGCAGAAACCAACAGATACGCTTTCACCTTGGATGACAGATTTT 1053

661 ATGCGCAGAAACCAACAGATACGCTTTCACCTTGGATGACAGATTTT 705

RESULT 9

AAC78111

ID AAC78111 standard; cDNA; 1989 BP.

XX AAC78111;

AC AAC78111;

XX

DT 08-FEB-2001 (first entry)

XX

DE Human cancer associated gene sequence SEQ ID NO:505.

XX

KW Human; cancer associated gene; cancer antigen; detection; cancer;

KW diagnosis; cytostatic; proliferative; cancer; immunomodulator;

KW antidiabetic; antihistaminic; antirheumatic; antiarthritic; antiviral;

KW antiinflammatory; antithyroid; antiallergic; antitubercular; cardiac;

KW dermatological; neuroprotective; thrombolytic; coagulant; nootropic;

KW vasotropic; antipsoriatic; antidiabetic; antihypertensive; inflammation;

KW immune disorder; haematopoietic cell disorder; autoimmune disorder;

KW allergic reaction; graft versus host disease; organ rejection;

KW haemostatic; thrombolytic; cardiovascular disorder; infection;

KW neurological disease; drug screening; ss.

XX

OS Homo sapiens.

XX

PN WO200055350-A1.

XX

PD 21-SEP-2000.

XX

PF 08-MAR-2000; 2000WO-US005882.

XX

PR 12-MAR-1999; 99US-0124270P.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

XX

PI Rosen CA, Ruben SM;

XX

XX WPI; 2000-587533/55.

DR P-PSDB; AAB43902.

XX

PT Novel isolated nucleic acids comprising sequences encoding peptides

PT useful for treating or diagnosing e.g. cancer.

XX

PS Claim 1; Page 1035-1036; 2352pp; English.

XX

CC AAC77607 to AAC78448 encode the human cancer associated proteins given in

CC AAB43398 to AAB44239. The proteins can have activities based on the

CC tissues and cells the genes are expressed in. Example of activities

CC include: cytostatic; proliferative; cancer; immunomodulator;

CC anti-diabetic; antiasthmatic; antirheumatic; antiarthritic;
 CC anti-inflammatory; antihypertensive; antiallergic; antibacterial; antiviral;
 CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
 CC neoplastic; vasotropic; antipsoriatic and antiangiogenic. The
 CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of
 CC the present invention
 XX
 SQ Sequence 1989 BP; 525 A; 439 C; 476 G; 544 T; 0 U; 5 Other;
 Query Match 54.9%; Score 578.4; DB 3; Length 1989;
 Best Local Similarity 75.0%; Pred. No. 1.1e-154;
 Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;
 QY 89 ATGAAGAGCCATCTTAAGAAGAGTTCTGCGCGTTTGTCTCTTCCAAATCCAGTACC 148
 DB 280 AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTGTCTCTTCCCGATCCAGTACC 339
 QY 149 CTGATATTGGAAATGTAAACAGCAGCGCTTCTTCTGAGCAGCAGAGAGGTGG 208
 DB 340 ATGATATCTGGCAGATGTAAAGAGCAGAGCTTCTTCTGAGCGCGAGAGGTGG 399
 QY 209 ACTTATCAAGAGATCTCCCTCACTGGAACCAAGCTTAAAGCAGATGAGAAGTACTTCACT 268
 DB 400 ACCTCTCCAGGACATCTGACCTGGATCCCTGAAACCGAGGAGAGATATTTATAT 459
 QY 269 CTCACATCTAGCTTTTGTGAGCAGGTGATGAATGTAAATGAAATTTGGTGAGC 328
 DB 460 CCATGTTCTGGCTTCTTGTGAGCAGCGATGGCATAGTAAATGAAATCTTGGTGAGC 519
 QY 329 GCTTTAGTCAGAGGTGAGGTTCAGAGCGCTGCTTCTTCTATGCTTCAATTTCTCA 388
 DB 520 GATTTAGCCAGAGGTTCAGATACAGAGCGCGCTTCTTCTATGCTTCCAAATTTGCCA 579
 QY 389 TCGAGAGATGTCTCAGAGATGTACAGTTTGTCTGATAGACATCTACATCAGAGATCCCA 448
 DB 580 TGAAGAACATACATCTGAAATGTATAGTCTTCTTATGACATCTACATAAAGATCCCA 639
 QY 449 AGAAAGGGAAATTTTATTTAATGCAATTTGAACACCATGCGCTATGTTAAGAAAAAGCAG 508
 DB 640 AAGAAAGGAAATTTCTTCAATGCCATTTGAACGATGCTTGTGTCAGAGAGAGCAG 699
 QY 509 ATTGGGCTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 568
 DB 700 ACTGGGCTTGGCTGATTTGGGGAACAAAGAGGCTACCTATGCTGAACGTTGTAGCCT 759
 QY 569 TTGCTGCTGATGAGAGGTTTCTTCTCAGATCTTCTGCTGCTATATTTCTGCTAAGA 628
 DB 760 TTGCTGCTGATGAGGCAATTTCTTCTGCTGCTTCTTCTGCTGCTATATTTCTGCTCAAGA 819
 QY 629 AGAGAGGTCTTATGCGCAGGACTCACCTTTTCCAAATGAACTCATCAGCAGAGATGAAGGAC 688
 DB 820 AACGAGGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 879
 QY 689 TTCATGTCATTTGCTTCCCTGATGTTTCAATTTCTAGTAAATAGCTTCAAGAGAA 748
 DB 880 TACACTGTGATTTGCTTCCCTGATGTTCAACACACCTGGTACACAAACCATCGAGGAGA 939
 QY 749 GGGTCAGGAGATCATTTGATGCTGCTCAATTTGAGCAGGAGGTTTAAACAGAGCCT 808
 DB 940 GAGTAAAGAGAAATATTTCAATGCTGTTTCGATAGAACAGGAGGTTTCTTCACTGAGCCT 999
 QY 809 TGCCAGTTGGCCTCATTTGGAATGAAATTTGCAATTTTGATGAAACAGTACATTTGAGTTGTAG 868

DB 1000 TGCTGTGAAGCTATTTGGGATGATTTGCATCTTAATGAGCAATACATTTGATTTGCG 1059
 QY 869 CTGACAGATTACTTTGTGGAACCTTGGATTTCTCAAGGTTTTCAGGCAAGAAATCCTTTTG 928
 DB 1060 CAGACAGACTTATGCTGGAACCTTGGTTTTCAGAGTTTTCAGAGTAGAGAACCCATTG 1119
 QY 929 ATTTTATGGAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACAGATTTTCAG 988
 DB 1120 ACTTTATGGAGATATTTCTTCTGGAAGGAAAGACTTAACCTTCTTTGAGAGAGATGGCG 1179
 QY 989 AGTATCAGCGTTTTCAGTTATGTCAGAAACCCAGAGATAACGCTTCTTACCTTTGGATGAC 1048
 DB 1180 AGTATCAGAGATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1239
 QY 1049 ATTT 1052
 DB 1240 ACTT 1243
 RESULT 10
 ADK70302
 ID ADK70302 standard; cDNA; 2215 BP.
 XX ADK70302;
 DT 06-MAY-2004 (first entry)
 XX Respiratory disease differentially expressed cDNA #38.
 DE
 XX ds; gene; cytostatic; respiratory; antiasthmatic; gene therapy;
 KW differential gene expression; respiratory disorder; lung cancer;
 KW chronic obstructive pulmonary disease; emphysema; asthma.
 XX Homo sapiens.
 XX WO2003101283-A2.
 XX 11-DEC-2003.
 XX 02-JUN-2003; 2003WO-US017409.
 XX 04-JUN-2002; 2002US-0386005P.
 XX (INCY-) INCYTE CORP.
 XX Rickert PK, Krasnow R;
 XX WPI; 2004-042945/04.
 XX New combination comprising cDNAs and proteins that are differentially
 XX expressed in respiratory disorders, useful for diagnosing or treating
 XX respiratory diseases e.g. lung cancer, chronic obstructive pulmonary
 XX diseases or asthma.
 XX Claim 1; SEQ ID NO 38; 343pp; English.
 XX The invention relates to cDNA sequences that are differentially expressed
 XX in respiratory disorders or their complements or encoded proteins. The
 XX cDNAs and proteins are useful for diagnosing, treating or monitoring
 XX treatment of a subject with a respiratory disease including lung cancer,
 XX chronic obstructive pulmonary diseases, emphysema or asthma. The protein
 XX is also useful for screening molecules or compounds to identify at least
 XX one ligand which specifically binds the protein. It is also useful for
 XX preparing and purifying a polyclonal or monoclonal antibody. This
 XX sequence corresponds to a cDNA of the invention.
 SQ Sequence 2215 BP; 593 A; 476 C; 523 G; 623 T; 0 U; 0 Other;
 Query Match 54.9%; Score 578.4; DB 12; Length 2215;
 Best Local Similarity 75.0%; Pred. No. 1.2e-154;
 Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

Db 464 CCATGTTCTGGCTTTCTTTGACGACGAGTGCATAGTAATGAAACTTGTGTGAGC 523
 Qy 329 GCTTTAGTCAGGAGGTCAGGTTCCAGAGGCTGCTGTTCTATGCTTTCAATTTCTCA 388
 Db 524 GATTAGCAAGAAGTTCAGATTACAGAAGCCGCTGTTCTATGCTTTCAATTTCTCA 583
 Qy 389 TCAGAAATGTTCACTCAGAGATGTACAGTTTGTGTGATAGACACTTACATCAGAGATCCCA 448
 Db 584 TGGAAACATACATCTCGAAATGTATAGTCTTCTTATTGACACTTACATAAAGATCCCA 643
 Qy 449 AGAAAGGGAATTTTATTAATGAATTCGAATTCGAACACCTGCTATGTAAGAAAGACAG 508
 Db 644 AGAAAGGGAATTTCTCTCAATGCGCAATGAACGATGCGCTTGTGTCAAGAAGACAG 703
 Qy 509 ATTGGGCTTTCGATGATAGACAGATAGAAATCTACTTTGGGAAAGAGTGTGGCCT 568
 Db 704 ACTGGGCTTTCGCTGATGAGGAGCAAGAGGCTACTATGATGAGCTGTTGTAGCCT 763
 Qy 569 TTCTGCTGTAGAGGAGTTTCTTCTCAGGATCTTTTGTCTATATTTCTGCTAAAGA 628
 Db 764 TTCTGCTGAGTGAAGGCAATTTCTTTCCGGTCTTTTGGTCTGATATTTCTGGCTCAAGA 823
 Qy 629 AGAGAGTCTTATGCCAGGACTCATTCTTTTCCAAATGAACTATCAGCAGAGATGAAGGAC 688
 Db 824 AACGAGGACTGATGCGCTGCTCACAATTTTCTAATGAACCTTATTAGCAGAGATGAGGTT 883
 Qy 689 TTCACTGTGACTTTGCTGCTGATGTTCCAAATCTAGTAATAGCCCTTCAGAGAAA 748
 Db 884 TACACTGTGATTTGCTGCTGATGTTCAACACCTGTTACAAACCATCGGAGGAGA 943
 Qy 749 GGTTCAGGAGATCATTTGTTGCTGCTCAAAATTCAGCAGGAGTTTAAACAGAGCCT 808
 Db 944 GAGTAGAGAAATTAATCAATGCTGTTCCGATAGAACAGGAGTTCTCTACTGAGCCT 1003
 Qy 809 TGCAGTGTGCTCATGTGAATGAATTTGATTTGATGAACAGTACATTTGAGTTGTAG 868
 Db 1004 TGCTCTGAAGCTCAATGGGATGAATGCACTCTAATGAAGCAATACATTTGAGTTGTG 1063
 Qy 869 CTGACAGATTTACTTGGGACTTGGATTTCTCAAGTTTTCAGGAGGAAATTCCTTTG 928
 Db 1064 CAGACAGACTTATGCTGGAAGTGGGTTTATGCAAGTTTTCAGAGTAGAGAACCCATTG 1123
 Qy 929 ATTTATGGAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACGAGTTTCAG 988
 Db 1124 ACTTTATGAGATATTTCTGGAAGAAAGTAACTTCTTTGAGAGAGATAGGCG 1183
 Qy 989 AGTATCAGCGTTTTCAGTTATGGCAGAAACCAACAGATACGTTCTTCACTTGGATGCG 1048
 Db 1184 AGTATCAGAGGATGGGAGTGTATGTCAGTCAAGTCCACAGAGAAATCTTTTACCTTGGATGCTG 1243
 Qy 1049 ATTT 1052
 Db 1244 ACTT 1247

RESULT 12
 AAS44917/c
 ID AAS44917 standard; DNA; 2482 BP.
 AC AAS44917;
 XX
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Human contig polynucleotide sequence #170.
 XX
 KW Mammal; human; rhesus monkey; baker's yeast; fission yeast; Norway rat;
 KW mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;
 KW cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;
 KW nervous system disorder; inflammatory disorder; cell differentiation; db;
 KW angiogenesis; stem cell growth factor; activin; inhibin; cartilage; burn;
 KW genetic disorder; bone regeneration; tendon; ligament; tissue repair;
 KW cytostatic; antirheumatic; antiarthritic; vulnery; antiinflammatory;

KW antibacterial; immunosuppressive; vasotropic; antiparkinsonian;
 KW neuroprotective; osteopathic; antidiabetic; antiasthmatic; antiallergic;
 KW immunostimulant; analgesic; gene therapy.
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN WO200164834-A2.
 XX
 PD 07-SEP-2001.
 XX
 XX 26-FEB-2001; 2001WO-US004926.
 XX
 XX 28-FEB-2000; 2000US-00515126.
 PR 18-MAY-2000; 2000US-00577409.
 PR 17-JUN-2000; 2000US-00597707.
 PR 14-JUL-2000; 2000US-00616807.
 PR 19-SEP-2000; 2000US-00664641.
 XX
 (HYSE-) HYSEQ INC.
 PA
 PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;
 PI Drmanac R;
 DR
 DR WPI; 2001-589862/66.
 DR P-PSDB; AAU28017.
 XX
 XX Novel polypeptides and nucleic acids obtained from cDNA libraries
 PT prepared from various human tissues, for diagnosis, treatment of cancer,
 PT neurological, inflammatory disorders and for use in arrays for detection.
 PS
 PS Claim 1; SEQ ID NO 514; 153pp; English.
 XX
 CC Sequences AAS44576-AAS44919 represent full-length polynucleotides and
 CC contig polynucleotides encoding polypeptides of the invention. The DNA
 CC and protein sequences are useful for the treatment, diagnosis and
 CC prevention of various types of disorder in a mammalian subject such as a
 CC human, dog, monkey, mouse, hamster or rat. The disorders include cancers
 CC such as leukaemia, lymphoma and neuroblastoma, autoimmune disorders such
 CC as multiple sclerosis, connective tissue disease, rheumatoid arthritis,
 CC diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system
 CC disorders such as Parkinson's disease, Alzheimer's disease, Huntington's
 CC chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and
 CC Wernicke disease, inflammatory disorders such as nephritis, Crohn's
 CC disease, ischaemia-perfusion injury, shock, sepsis and inflammatory
 CC bowel disease. The sequences exhibit activity relating to angiogenesis,
 CC cell proliferation, cell differentiation, stem cell growth factor,
 CC activin or inhibin. Therefore, they can be used to manipulate stem cells
 CC in culture to give rise to neuroepithelial cells that can be used to
 CC augment or replace cells damaged by illness, accidental damage or genetic
 CC disorders. The sequences may also be used for regeneration of bone,
 CC cartilage, tendons and ligaments and in tissue repair and burn healing.
 CC Note: Some sequences for this patent did not form part of the printed
 CC specification, but were obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 2482 BP; 707 A; 598 C; 522 G; 655 T; 0 U; 0 Other;
 Query Match 54.9%; Score 578.4; DB 4; Length 2482;
 Best Local Similarity 75.0%; Pred. No. 1.2e-154;
 Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;
 Qy 89 ATGAAGAGCCACTCTTAAGAAAGAGTTCTCGCGGTTTGTCACTTTTCAATCCAGTACC 148
 Db 2086 AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTTGTCTATCTTCCCATCGAGTACC 2027
 Qy 149 CTGATATTTGGAATATGTAAGACAGCAGCAGGCTTCTTCTGGACACAGAGAGTCCG 208
 Db 2026 ATGATATCTGGCAGATGTAAGAAAGGAGGAGGCTTCTTTTGGACCCCGAGGAGTTG 1967
 Qy 209 ACTTATCAAGAGGATCTCCCTCCTCCTGGAACCAAGCTTAAAGACAGATCAGAGTACTTCACT 268

PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0244864P.
XX
PA (AVAL-) AVALON PHARM.
XX
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
XX WPI; 2002-188264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX
XX Claim 1; SEQ ID NO 4196; 44pp; English.
XX
XX The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC tumour
XX
XX Sequence 2500 BP; 679 A; 522 C; 595 G; 704 T; 0 U; 0 Other;
XX
XX Query Match 54.9%; Score 578.4; DB 6; Length 2500;
XX Best Local Similarity 75.0%; Pred. NO. 1.2e-154;
XX Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;
XX
XX 89 ATGAGAGCCACTCTTGAAGAGAGTTCTGCGGGTTTGTATCTTTCCATCCAGTACC 148
XX
XX 397 AGGATGAGCGCTGTCTGAGAGAAACCCCGCGCTTTGTCTATCTTCCCATCGAGTACC 456
XX
XX 149 CTGATATTGGAAATGTATAACAGCGCAGCGCTTCTCTCTGGACAGCAGAGAGGTGC 208

Db 457 ATGATATCTGCAGATGTATAAGAGCGAGAGGCTTCTTTTGGACCCCGAGGAGTTG 516
Qy 209 ACTTATCAAAGGATCTCCCTCACTCGGAACAAGCTTTAAAGCAGATGAGAAGTACTTTCATCT 268
Db 517 ACCTCTCAAGGACATTCAGCAGCTGGGATCCCTGAAACCCGAGGAGAGATATTTATAT 576
Qy 269 CTCACATCTTAGCCTTTTTCAGCCAGTGAAGAAATGTAAATGAAATTTGGTGGAGC 328
Db 577 CCCATGTTCTGGCTTTCTTTGACAGCAAGCGATGGCATAGTAAATGAAAACTTGGTGGAGC 636
Qy 329 GCTTTAGTCAGAGGTGCGGTTCCAGAGGCTCGCTGTTTCTATGGCTTCAAAATCTCA 388
Db 637 GATTTAGCCAAGAGTTTCAAGATTACAGAAGCCGCTGTTTCTATGGCTTCCAATTTGCCA 696
Qy 389 TCGAGAAATGTTCACTCAGAGATGTACAGTTTGTGATAGACACTTACATCAGAGATCCCA 448
Db 697 TGGAAACATACATTTCTGAATGTATAGTCTCTTTATTTGACACTTACATAAAGATCCCA 756
Qy 449 AGAAAGGGAATTTTATTTAATGCAATTTGAAACCATGCCCTATGTTAAGAAAAAGCAG 508
Db 757 AAGAAAGGGAATTTCTCTTCAATGCCATTTGAAACGATGCCCTTGTCTCAAGAAAGAGCAG 816
Qy 509 ATTGGGCTTTCGATGATAGCAGATAGAAATCTACTTTTGGGGAAGAGTGGTGGCT 568
Db 817 ACTGGGCTTTCGCTGGATTTGGGCAACAGAGGCTACCTATGTTGAACGTGTTGTAGCCT 876
Qy 569 TTGCTGCTGTAGAGGAGTTTCTTCTCAGGATCTTTTCTGCTGCTATATTTGCGCTAAAGA 628
Db 877 TTGCTGAGTGAAGAGGCAATTTCTTTCCGGTCTTTTTCGGTTCGATATTTCTGAGCTCAAGA 936
Qy 629 AGAGAGGTCTTATGCCAGGACTCACTTTTTCATGAACTCATCAGCAGAGATGAAGGAC 688
Db 937 AACGAGAGCTGATGCCCTGGCTCACATTTTCTATGAACTTATAGCAGAGATGAGGTT 996
Qy 689 TTCACTGTGACTTTGCTTGGCTGATGTTTCCAAATCTTAGTAAATAGGCTTTCAGAAAGAAA 748
Db 997 TACACTGTGATTTTGTCTGCTGATGTTTCAAAACCTGCTGACAAACCATCGAGAGGAGA 1056
Qy 749 GGGTCAGGAGATCATTTGATGCTGTCAAAATTTGACAGAGGTTTTCAGAGAGCCT 808
Db 1057 GAGTAAAGAGAAATAAATTTATCAATGCTGTTCGGATAGAAACAGGAGTTCTCTCAGGCT 1116
Qy 809 TGCAGTTGGCTTCATTTGGAATGAATTTGATGAAACAGTACATTTGAGTTTGTAG 868
Db 1117 TGCTGTGAAGCTCATTTGGGATGAATTTGCACTCTAATGAGCAATACATTTGATTTGTGG 1176
Qy 869 CTGACAGATTACTTTGTGGAACCTTGGATTTCTCAAAGGTTTTCAGGCAGAAATTCCTTTG 928
Db 1177 CAGACAGACTTATGCTGGAAGTGGGTTTTCAGAGGTTTTCAGAGTAGAGAACCCATTG 1236
Qy 929 ATTTTATGGAACATTTCTTTTGAAGAAACAAATTTCTTTTGAAGAAACAGATTTCAG 988
Db 1237 ACTTTATGGAATATTTTCACTTGAAGGAAAGACTAATCTCTTTTGAAGAGAGATGAGCG 1296
Qy 989 AGTATCAGCCTTTTTCAGGTTTATGSCAGAAACCCAGATAACGCTTTCACCTTGGATGAG 1048
Db 1297 AGTATCAGAGATGGGAGTGTCAAGTCCACAGAGAAATCTTTTACCTTTGGAATGCTG 1356
Qy 1049 ATTT 1052
Db 1357 ACTT 1360

Search completed: June 16, 2005, 07:10:00
Job time : 677 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 16, 2005, 09:39:47 ; Search time 640 Seconds
(without alignments)
3246.608 Million cell updates/sec

Title: US-10-698-228-1
Perfect score: 1821
Sequence: 1 MGDPEPEAAGLDQDERSSS.....QRFAYMAETTDNVFTLDADF 351

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlh
-Q=/cgn2_1/USFTO_spool/US10698228/runat_14062005_161415_28287/app_query.fasta_1.519
-DB=N Geneseq 16dec04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10698228 @CNG 1 1 644 @runat_14062005_161415_28287 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq_16Dec04.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1821	100.0	1053	4	Aaf32438 Human rib
2	1821	100.0	1053	4	Aaf32447 Human rib
3	1821	100.0	1081	4	Aaf32440 Human rib
4	1821	100.0	2596	3	Aa12411 cDNA enco
5	1821	100.0	4955	4	Aaf32439 Human rib

6	1821	100.0	4955	13	ADR24210
7	1817	99.8	1601	4	AaH14924
8	1460.5	80.2	1989	3	AAC78111
9	1460.5	80.2	2215	12	ADK70302
10	1460.5	80.2	2216	10	ADJ56536
11	1460.5	80.2	2482	4	AS44917
12	1460.5	80.2	2500	6	ABL65414
13	1460.5	80.2	2500	6	ABL65517
14	1460.5	80.2	2500	6	ABL65859
15	1460.5	80.2	2500	8	ABX10335
16	1460.5	80.2	2500	11	ADI32044
17	1460.5	80.2	2500	12	ADN04443
18	1460.5	80.2	2500	12	ADN04443
19	1460.5	80.2	2500	13	ADQ09273
20	1460.5	80.2	3294	4	AAH73225
21	1460.5	80.2	3393	12	ADN03788
22	1460.5	80.2	3393	12	ADO19225
23	1447.5	79.5	2641	4	AS44745
24	1429	78.5	1170	2	AAV05641
25	1425.5	78.3	1328	12	ADO57308
26	1376.5	75.6	1371	5	AA579474
27	1348	74.0	1289	4	ABL14627
28	1241.5	68.2	3945	4	ABL14626
29	1230	67.5	1218	8	ABT20895
30	1220.5	67.0	1292	13	ADS49379
31	1210	66.4	14176	2	AAT84564
32	1210	66.4	14176	4	AAP84949
33	1210	66.4	14176	4	AAP84948
34	1192.5	65.5	1146	8	ABT19075
35	1192	65.5	1242	6	ABZ32250
36	1188	65.2	706	13	ADQ57092
37	1176.5	64.6	963	13	ADS58378
38	1168.5	64.2	1200	10	ACC61238
39	1168.5	64.2	1200	10	ADK63241
40	1168.5	64.2	1200	13	ADT47419
41	1161.5	63.8	1173	13	ADT47814
42	1161.5	63.8	1206	10	ADB69740
43	1151	63.2	1314	8	ABT18481
44	1151	63.2	1450	8	ABT20297
45	1151	63.2	3314	8	ABT17887

ALIGNMENTS

RESULT 1
AAF32438
ID AAF32438 standard; cDNA; 1053 BP.
AC AAF32438;
XX
XX
DT 18-APR-2001 (first entry)
XX
DE Human ribonucleotide reductase TP53R2H nucleotide sequence SEQ ID NO:2.
XX
KW Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.
XX
OS Homo sapiens.
XX
FN WO200100799-A1.
XX
PD 04-JAN-2001.
XX
PF 27-JUN-2000; 2000WO-JP004189.
XX
PR 28-JUN-1999; 99JP-00181131.
PR 06-JUL-1999; 99JP-00192391.
PR 21-JAN-2000; 2000JP-00017770.
XX
XX (TAKE) TAKEDA CHEM IND LTD.
XX (NAKA/) NAKAMURA Y.
XX Nakamura Y, Arakawa H, Tanaka H;

DR WPI: 2001-112446/12.
 DR P-PSDB: AAB69050.

PT Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
 PT diagnosis, treatment and prevention of cancer.

PS Claim 5; Fig 1-3; 102pp; Japanese.

XX The present sequence encodes a human ribonucleotide reductase designated
 CC TP53R2H. TP53R2H has cytosstatic activity. TP53R2H is part of the DNA
 CC repair mechanism and its activity is induced by p53. It can be used for
 CC the treatment, prevention and diagnosis of a wide range of cancers

XX Sequence 1053 BP; 314 A; 192 C; 246 G; 302 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 4,17e-212 Length: 1053
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-10-698-228-1 (1-351) x AAF32438 (1-1053)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20

DB 1 ATGGGCGACCCGGAAGGCGGAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60

QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40

DB 61 GACACCAACGAAGTGAATAAGTCAATGAAGCCACTCTCTAAGAAAGAGTTCTCGC 120

QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60

DB 121 CGGTTTGTCATCTTCCATCCAGTACCTCTGATATTGGAAATGATATAACAGGCACAG 180

QY 61 AlaSerPheTyrThrAlaGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80

DB 181 GCTTCTCTGGACAGCAGAGAGGTGCGACTTATCAAGAGTCTCCCTCAGTGAACAAG 240

QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaIleSerAsp 100

DB 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCAGATCTTACGCTTTTTCGACGAGTAT 300

QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120

DB 301 GGAATTGTAAATGAATTTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCT 360

QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140

DB 361 CGCTGTTTCTATGGCTTTCAATTTCTCATCGAATGTTTCACTCAGAGATGTACAGTTG 420

QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160

DB 421 CTGATAGACACTTACATCAGATATCCAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 480

QY 161 ThrMetProTyrValLysLysAlaAspTyrAlaLeuArgTyrIleAlaAspArgLys 180

DB 481 ACCATGCCCTATGTTAAGAAAGAGAGATTTGGGCTTGGATGGATGATAGAGAA 540

QY 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200

DB 541 TCTACTTTTGGGAAAGAGTGGTGGCTTGTGCTGTAGAGAGATTTTCTCTCAGGA 600

QY 201 SerPheAlaAlaIlePheTyrLysLysArgGlyLeuMetProGlyLeuThrPheSer 220

DB 601 TCTTTTGTCTATATTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 660

QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240

DB 661 AATGAACATCATCAGCAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720

QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260

DB 721 TACTTAGTAATAAGCCTTCAGAAAGAGGTCAGGAGAGATCATTTGTGCTGTCAAA 780

QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280

DB 781 ATTGACGACGAGTTTAAACAGAAAGCCTTGCAGTTCGCTCATTTGAATGAATTCATT 840

QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300

DB 841 TTGATGAACAGTACATTTGATTTGTAGCTGCACAGATTACTTTGTGAACCTTGGATTTCA 900

QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320

DB 901 AAGTTTTTTCAGGACAGAAATCTTTTGTATTTATGAAACATTTCTTTTGAAGGAAAA 960

QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340

DB 961 ACAATTTCTTTGAGAAACGAGTTTTCAGATATCAGGTTTTCAGATTATGCGAGAAACC 1020

QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

DB 1021 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1053

RESULT 2

AAF32447

ID AAF32447 standard; cDNA; 1053 BP.

XX AAF32447;

DT 18-APR-2001 (first entry)

XX Human ribonucleotide reductase related nucleotide sequence SEQ ID NO:12.

XX Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.

XX Homo sapiens.

PN WO200100799-A1.

PD 04-JAN-2001.

XX 27-JUN-2000; 2000WO-JP004189.

XX 28-JUN-1999; 99JP-00181131.

PR 06-JUL-1999; 99JP-00192391.

PR 21-JAN-2000; 2000JP-00017770.

XX (TAKE) TAKEDA CHEM IND LTD.

PA (NAKA/) NAKAMURA Y.

XX Nakamura Y, Arakawa H, Tanaka H;

XX WPI; 2001-112446/12.

XX Ribonucleotide reductase involved in DNA repair and DNA encoding it, for

XX diagnosis, treatment and prevention of cancer.

XX Claim 6; Page 96-97; 102pp; Japanese.

XX The present invention describes a human ribonucleotide reductase

XX designated TP53R2H. TP53R2H has cytosstatic activity. TP53R2H is part of

XX the DNA repair mechanism and its activity is induced by p53. It can be

XX used for the treatment, prevention and diagnosis of a wide range of

XX cancers. The present sequence represents a human ribonucleotide reductase

XX related sequence which is given in the exemplification of the present

XX invention

SQ Sequence 1053 BP; 314 A; 191 C; 246 G; 302 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4.17e-212 Length: 1053

Score: 1821.00 Matches: 351

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-698-228-1 (1-351) x AAF32447 (1-1053)

```
QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20
DB 1 ATGGCGCAGCCGGAAGGCGGAAGCGCGGCGCTGGATCAGGATGAGAGATCATCTTCA 60
QY 21 AspThrAsnGluSerGluLeuIleYsSerAsnGluGluProLeuLeuArgLysSerSerArg 40
DB 61 GACACCAAGAAAGTAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 120
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
DB 121 CGGTTTGTCATCTTTCATCTCCAGTACCTGATATTTGGAAATGATATAACAGGACAG 180
QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
DB 181 GCTTCCTCTGGACAGCAGAGAGGTTGACTTATCAAGGATCTCCTCACTGGAACAAG 240
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaIaSerAsp 100
DB 241 CTTAAAGCAGATGAGAAGTACTTCTCTCATCTTACCTTTTGGCAGCCAGTGAT 300
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 301 GGAATTTGAAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 361 CGCTGTTCTATGGCTTTCAATCTCATCGAGATGTTCACTCAGAGATGACAGTTTG 420
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTGAA 480
QY 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
DB 481 ACCATGCCCTATGTTAAAGAAAAGCAGATTTGGCCCTTCGATGGATAGCAGATAGAAA 540
QY 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
DB 541 TCTACTTTTGGGAAAGATGGTGGCTTTGCTGCTGTAGAGAGATTTCTTCTCAGGA 600
QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 601 TCTTTTGTCTATATCTTGCTTAAAGAGAGAGGCTTTATGCGCAGGACTCACTTTTTC 660
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
DB 661 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
DB 721 TACTTAGTAAATGAGCTTTCAGAAAGAGGCTCAGGAGATCAATTTGATGCTGTCAAA 780
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
DB 781 ATTGACGAGAGTATTTTAAACAGAGGCTTGCCAGTGGCTCATTTGGAATGATTCAT 840
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
DB 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTACTTTGTGAACTTGGATTC 900
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
DB 901 AAGGTTTTTCAGGCGAATAATCTTTTGTATTTTATGGAATAACATTTCTTTAGAAGAAA 960
QY 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
```

```
DB 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAC 1020
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB 1021 ACAGATAACGTTCTTTCACCTTGGATGCAGATTTT 1053
```

RESULT 3

AAF32440
ID AAF32440 standard; cDNA; 1081 BP.

XX AAF32440;

XX 18-APR-2001 (first entry)

DE Human ribonucleotide reductase related nucleotide sequence SEQ ID NO:4.

XX Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.

XX Homo sapiens.

PN WO200100799-A1.

PD 04-JAN-2001.

XX 27-JUN-2000; 2000WO-JP004189.

XX 28-JUN-1999; 99JP-00181131.

PR 06-JUL-1999; 99JP-00192391.

PR 21-JAN-2000; 2000JP-00017770.

XX (TAKE) TAKEDA CHEM IND LTD.

PA (NAKA/) NAKAMURA Y.

XX Nakamura Y, Arakawa H, Tanaka H;

XX WPI; 2001-112446/12.

XX Ribonucleotide reductase involved in DNA repair and DNA encoding it, for

XX diagnosis, treatment and prevention of cancer.

XX Example 2; Page 91; 102pp; Japanese.

XX The present invention describes a human ribonucleotide reductase
CC designated TP53R2H. TP53R2H has cytostatic activity. TP53R2H is part of
CC the DNA repair mechanism and its activity is induced by p53. It can be
CC used for the treatment, prevention and diagnosis of a wide range of
CC cancers. The present sequence represents a human ribonucleotide reductase
CC related sequence which is used in an example from the present invention

XX SQ Sequence 1081 BP; 322 A; 199 C; 252 G; 308 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4.34e-212 Length: 1081
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-698-228-1 (1-351) x AAF32440 (1-1081)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20

DB 20 ATGGCGCAGCCGGAAGGCGGAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 79

QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40

DB 80 GACACCAAGAAAGTGAATTAAGTCAATGAGAGCCACTCTTAAGAAAGAGTTCTCG 139

QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60

DB 140 CGGTTTGTCTATCTTTCATTCAGTACCTCGATATTTGGAAATGTAATAACAGGCACAG 199

QY 61 AlaSerPheTrpThrAlaGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
 Db 200 GCTTCCTCTGACAGCAGAGAGGTGACTTATCAAGGATCTCCCTCACTGGAAACAAG 259
 QY 81 LeuLysAlaAspGluLysTrpPheHisLeuLeuAlaPhePheAlaAlaSerAsp 100
 Db 260 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTTTAGCCCTTTTTCAGCCAGTGAT 319
 QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 Db 320 CGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 379
 QY 121 ArgCysPheTrpGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTrpSerLeu 140
 Db 380 CGCTGTTTCTATGGCTTCAAAATCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG 439
 QY 141 LeuIleAspThrTrpIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 Db 440 CTGATAGACACTTACATCAGATCCCAAGAAAGGGAATTTTATTTAATGCATTTGAA 499
 QY 161 ThrMetProTrpValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
 Db 500 ACCATGCCCTATGTAAGAAAAGCAGATTGGCCCTTCGGATGGATAGCAGATAGAAA 559
 QY 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
 Db 560 TCTACTTTTGGGCAAGAGTGGTGGCTTTCGTCTGTAGAAAGGAGCTTTCTTCTCAGGA 619
 QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTrpPheSer 220
 Db 620 TCTTTTGTGCTATATTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 679
 QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 Db 680 ATGAACTCATCAGCAGATGAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 739
 QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
 Db 740 TACTTAGTAAATAAGCCTTCAGAAAGAGGTGAGGAGATCATTTGATGCTGTCAAA 799
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 Db 800 ATTGAGCAGAGTGTATTAACAAGAGCTTGGCCAGTTGGCTCATTTGGAATGAAATTCAT 859
 QY 281 LeuMetLysGlnTrpIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 Db 860 TTGATGAACAGTACATTGAGTTGTAGCTGACAGATTACTTGTGAACTTGGATTTCTCA 919
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 Db 920 AAGGTTTTTTCAGGCAGAAAATCCTTTTGTATTTATGAAACATTTCTTTAGAGGAAAA 979
 QY 321 ThrAsnPhePheGluLysArgValSerGluTrpGlnArgPheAlaValMetAlaGluThr 340
 Db 980 ACAAAATTTCTTTGAGAAACAGTTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC 1039
 QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 Db 1040 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1072

RESULT 4

ID AAA12411 standard; cDNA; 2596 BP.

XX XX AAA12411;

XX XX 25-JUL-2000 (first entry)

XX XX cDNA encoding a human RNA-associated protein.

DE Human; RNA-associated protein; cell proliferation; cancer; inflammation;
 XX immune response; reproductive disorder; actinic keratosis;
 KW atherosclerosis; arteriosclerosis; bursitis; cirrhosis; hepatitis;
 KW

XX mixed connective tissue disease; myelofibrosis; primary thrombocythemia;
 KW paroxysmal nocturnal hemoglobinuria; polycythemia vera; psoriasis;
 KW trauma; ss.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FT CDS 56..1111
 FT FT /*tag= a
 XX /product= "RNA-associated protein"
 PN W0200015799-A2.
 XX 23-MAR-2000.
 XX 17-SEP-1999; 99WO-US021688.
 XX 17-SEP-1998; 98US-00156039.
 PR 22-SEP-1998; 98US-00158720.
 PR 04-NOV-1998; 98US-00186815.
 PR 08-APR-1999; 99US-0128660P.
 XX (INCY-) INCYTE PHARM INC.
 XX Tang YT, Corley NC, Guegler KJ, Gorgone GA, Patterson C;
 PI Hillman JL, Baughn MR, Lal P, Azimzai Y, Yue H, Yang J;
 XX WPI: 2000-271437/23.
 DR P-PSDB; AAY84439.
 XX New polypeptides and polynucleotides, useful for preventing and treating
 PT a disorder associated with increased or decreased expression of RNA
 PT associated proteins.
 XX Claim 9; Page 120-121; 131pp; English.

CC The present sequence encodes a human RNA-associated protein. The
 CC expression of RNA-associated proteins is closely associated with
 CC reproductive tissues, nervous tissues, cell proliferation including
 CC cancer, inflammation and immune responses, and so they may be used for
 CC diagnosis, treatment or prevention of cell proliferative, diseases and
 CC immune/inflammatory disorders, and reproductive disorders. Diseases and
 CC disorders which may be treated include actinic keratosis,
 CC atherosclerosis, arteriosclerosis, bursitis, cirrhosis, hepatitis, mixed
 CC connective tissue disease, myelofibrosis, paroxysmal nocturnal
 CC hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia
 CC and cancers, and trauma

SQ Sequence 2596 BP; 783 A; 418 C; 510 G; 885 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1-59e-211	Length:	2596
Score:	1821.00	Matches:	351
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	3	Gaps:	0

US-10-698-228-1 (1-351) x AAA12411 (1-2596)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnArgSerSerSer 20
 Db 56 ATGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 115
 QY 21 AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerSerArg 40
 Db 116 GACACCAACCAAGTGAATAAAGTCAATGAAGCCACTCTCTAAGAAAGAGTTCTCGC 175
 QY 41 ArgPheValIlePheProIleGlnTrpProAspIleTrpLysMetTrpLysGlnAlaGln 60
 Db 176 CGGTTTGTCTACTTTCATATCCAGTACCTGATATTGGAAAAATGTATAACAGGCACAG 235
 QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80

```

Db      236  GCTTCCTCTTGGACACAGAGAGGTCGACTTATCAAGGATCTCCCTCACTGGAAACAAG 295
Qy      81  LeuLeuAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaLaSerAsp 100
Db      296  CTTAAAGCAGATGAGAAGTACTTCACTCTCACATCTTAGCCTTTTGGCAGCAGTGAT 355
Qy      101  GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db      356  GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGCCT 415
Qy      121  ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db      416  CGCTGTTTCTATGGCTTTCAAAATTCATCGAAGTGTTCACCTCAGAGATGTACAGTTG 475
Qy      141  LeuIleAspThrTyrIleArgAspProLysLeuValArgGluPheLeuPheAsnAlaIleGlu 160
Db      476  CTGATAGACACTTACATCAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 535
Qy      161  ThrMetProTyrValLysLysAlaAspTyrAlaLeuArgTyrPheAlaAspArgLys 180
Db      536  ACCATCCCTATGTTAAGAAAAAAGCAGATTGGGCTTGGATGGATAGCAGATAGAAA 595
Qy      181  SerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePheSerGly 200
Db      596  TCTACTTTTGGGGAAAGAGTGGTGGCTTTGCTGTGTAGAGAGAGTTTCTTCTCAGGA 655
Qy      201  SerPheAlaAlaIlePheTyrLeuLysLeuValArgGlyLeuMetProGlyLeuThrPheSer 220
Db      656  TCTTTTGTCTATATCTGGCTTAAGAGAGAGGCTTATGCCAGGACTCACTTTTTC 715
Qy      221  AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db      716  AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA 775
Qy      241  TyrLeuValAsnLysProSerGluGluArgValArgGluIleLeuValAspAlaValLys 260
Db      776  TACTTAGTAATAAGCCTTCAGAAAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 835
Qy      261  IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db      836  ATTGACACAGAGTTTTTAACAGAAAGCCTTGCCAGTTGGCCTCATTTGGAATGCAAT 895
Qy      281  LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db      896  TTGATGAAACAGTACATTGAGTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTTCTCA 955
Qy      301  LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db      956  AAGGTTTTTTCAGGCAGAAAAATCCTTTTGTATTTATGMAAACATTTCTTTTAGAAGGAAA 1015
Qy      321  ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db      1016  ACAAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1075
Qy      341  ThrAsnValPheThrLeuAspAlaAspPhe 351
Db      1076  ACAGATAAGCTCTTCACCTTGGATGAGATTTT 1108

RESULT 5
AAF32439
ID  AAF32439 standard; cDNA; 4955 BP.
XX
AC  AAF32439;
XX
DT  18-APR-2001 (first entry)
XX
DE  Human ribonucleotide reductase related nucleotide sequence SEQ ID NO.3.
XX
KW  Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.
XX
OS  Homo sapiens.
XX

```

```

PN      WO200100799-A1.
XX      04-JAN-2001.
XX      27-JUN-2000; 2000WO-JP004189.
XX      28-JUN-1999; 99JP-00181131.
PR      06-JUL-1999; 99JP-00192391.
PR      21-JAN-2000; 2000JP-00017770.
XX      (TAKES ) TAKEDA CHEM IND LTD.
PA      (NAKA/) NAKAMURA Y.
XX      Nakamura Y, Arakawa H, Tanaka H;
XX      WPI; 2001-112446/12.
XX      Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
PT      diagnosis, treatment and prevention of cancer.
XX      Example 2; Page 87-90; 102pp; Japanese.
XX      The present invention describes a human ribonucleotide reductase
CC      designated TP53R2H. TP53R2H has cytosolic activity. TP53R2H is part of
CC      the DNA repair mechanism and its activity is induced by p53. It can be
CC      used for the treatment, prevention and diagnosis of a wide range of
CC      cancers. The present sequence represents a human ribonucleotide reductase
CC      related sequence which is used in an example from the present invention
XX      Sequence 4955 BP; 1526 A; 803 C; 999 G; 1627 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:      4.13e-211      Length:      4955
Score:          1821.00      Matches:      351
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:              4              Gaps:      0

US-10-698-228-1 (1-351) x AAF32439 (1-4955)
Qy      1  MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20
Db      245  ATGGCGCACCCGAAAGGCGCGGCGGCGGATCAGGATCAGAGATCATCTTCA 304
Qy      21  AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
Db      305  GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTCAAGAAAGAGTTCTCGC 364
Qy      41  ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
Db      365  CGGTTTGTATCTTTCCATCCAGTACCTCATATTTGGAATAATGTATTAACAGGCACAG 424
Qy      61  AlaSerPheThrThrAlaGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
Db      425  GCTTCCTTCTGGACAGAGAGGTGACTTATCAAGAGTCTCCCTCACTGGAACAAG 484
Qy      81  LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaLaSerAsp 100
Db      485  CTTAAAGCAGATGAGAAGTACTTCACTCTCACATCTTAGCCTTTTTCAGCAGCATGAT 544
Qy      101  GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db      545  GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGCCT 604
Qy      121  ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db      605  CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAAGTGTTCACCTCAGAGATGTACAGTTG 664
Qy      141  LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db      665  CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 724

```

QY 161 ThrMetProTyrVallylsLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
 Db 725 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGCCCTTGCATGGATAGCAGATAGAAAA 784
 QY 181 SerThrPheGluArgValValAlaPheAlaValGluGlyValPhePheSerGly 200
 Db 785 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTAGAGAGTTCCTTCTCAGGA 844
 QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 Db 845 TCTTTTGTCTATATCTTGGCTTAAAGAGAGAGGCTTATGTCAGGAGCTCACTTTTCC 904
 QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 Db 905 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTTGGCTTGGCTGATGTTCCAA 964
 QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAlaValLys 260
 Db 965 TACTTAGTAAATAGCCTTCAGNAGAAAGGTGAGGAGATCAATTGTTGATGCTGTCAAA 1024
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 Db 1025 ATTGAGCAGGAGTTTAAACAGAGGCTTGGCCAGTTGGCTCATTTGGAATGAATTGCAIT 1084
 QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 Db 1085 TTGATGAAACAGTACATTGAGTTTGTAGCTGCAGAGATTACTTTGTGAACTTGGATTCTCA 1144
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 Db 1145 AAGGTTTTTCAGCAGAGAAATCCTTTTGATTTTATGGAAACATTTCTTTAGAGGAAAA 1204
 QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 Db 1205 ACAATTTCTTTGAGBAAACGAGTTTCAGAGTATCAGCGTTTGCAGTTATGGCAGAAACC 1264
 QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 Db 1265 ACAGATAAGCTCTTCACTTGGATGCAGATTTT 1297

RESULT 6
 ADR24210
 ID ADR24210 standard; DNA; 4955 BP.
 XX
 AC ADR24210;
 XX
 DT 21-OCT-2004 (first entry)
 XX
 DE Breast cancer prognosis marker #71.
 XX
 KW ds; breast cancer; prognosis; gene expression; diagnosis.
 XX
 OS Homo sapiens.
 XX
 XX WO2004065545-A2.
 XX
 XX 05-AUG-2004.
 XX
 PF 15-JAN-2004; 2004WO-US0001100.
 XX
 PR 15-JAN-2003; 2003US-00342887.
 XX
 XX (ROSE-) ROSETTA INPHARMATICS LLC.
 PA (NECA-) NETHERLANDS CANCER INST.
 XX
 XX Van't Veer LJ, He Y;
 PI WPI; 2004-593473/57.
 XX
 DR
 XX
 XX
 XX
 PT Classifying a breast cancer patient according to prognosis comprises
 PT determining the similarity between the level of expression of each of
 PT five genes in a cell sample taken from patient, to control levels.
 XX

PS Disclosure; SEQ ID NO 71; 226pp; English.
 XX
 CC The invention relates to a method of classifying a breast cancer patient
 CC according to prognosis by determining the similarity between the level of
 CC expression of each of five genes for which markers are listed in the
 CC specification, in a cell sample taken from the breast cancer patient, to
 CC control levels of expression for each respective five genes to obtain a
 CC patient similarity value. The methods are useful for classifying a breast
 CC cancer patient according to prognosis. Kits and computer program products
 CC are useful for data analysis using the diagnostic, prognostic and
 CC statistical methods of the invention. This sequence corresponds to a
 CC marker used in the method of the invention.
 XX
 SQ Sequence 4955 BP; 1526 A; 803 C; 999 G; 1627 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 4.13e-211 Length: 4955
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 13 Gaps: 0
 US-10-698-228-1 (1-351) x ADR24210 (1-4955)
 QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
 Db 245 ATGGGCGACCGGAAAGCGCGGCGGTGGATCAGGATGAGAGATCATCTTCA 304
 QY 21 AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerSerArg 40
 Db 305 GACACCAACGAAAGTGAATAAAGTCAATCAAGAGCCACTCTTAAGAAAGATTTCTCC 364
 QY 41 ArgPheValIlePhePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
 Db 365 CGGTTTGTATCTTTCATCCAGTACCTGATTTTGGAAATGTATTAACAGGCACAG 424
 QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
 Db 425 GCTTCTCTTGGACAGCAGAGAGGTGACATTATCAAGAGGATCTCCCTCACTGGAACAAG 484
 QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAspAsp 100
 Db 485 CTTAAAGCAGATGAGAGATGACTTCTCTCTCACATCTTTAGCCCTTTTTCAGGCCAGTGAT 544
 QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 Db 545 GGAATTGTAATGAAATTTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 604
 QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 Db 605 CGCTGTTTCTATGCTTTCAAAATTTCTCATCGAGAATGTTTCACTCAGAGATGTACAGTTTG 664
 QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTAATGCAATTGAA 724
 QY 161 ThrMetProTyrVallylsLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
 Db 725 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTGGCATGGATGAGATAGAAAA 784
 QY 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
 Db 785 TCTACTTTTGGGAAAGAGTGGTGGCTTTCCTGCTGTAGAGAGATTTTCTTCTCAGGA 844
 QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 Db 845 TCTTTTGTCTATATCTTGGCTTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTCC 904
 QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 Db 905 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTTGGCTTGGCTGATGTTCCAA 964

Qy 241 TyrLeuValAsnLysProSerGluGluArgValArgGluLeuLeuValAspAlaValLys 260
 Db 965 TACTTAGTAATAAGCCCTTCAGAAAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAAA 1024
 Qy 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuLeuGlyMetAsnCysIle 280
 Db 1025 ATTGAGCAGAGGTTTTTAACAGAGCCTTCCAGTTGGCTCATTTGGAATGATTCATT 1084
 Qy 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 Db 1085 TTGATGAACAGTACATGATGTTGTAGCTGACAGATTACTTTGGGAACCTGGATTCTCA 1144
 Qy 301 LysValPheGlnAlaGluAenProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 Db 1145 AAGTTTTTCAGGCAGAAATCCCTTTGATTTTATGGAAACATTTCTTTAGAGAGGAAA 1204
 Qy 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 Db 1205 ACAATTTCTTTGAGAAACAGTTTCAGAGTATCAGGTTTTCAGATTATGGCAGAAACC 1264
 Qy 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 Db 1265 ACAGATAACCTCTTCACCTTGGATGCAGATTTT 1297

RESULT 7

AAH14924
 ID AAH14924 standard; cDNA; 1601 BP.
 XX AC AAH14924;
 XX DT 26-JUN-2001 (first entry)
 XX DE Human cDNA sequence SEQ ID NO:12810.
 XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 XX OS Homo sapiens.
 XX PN EF1074617-A2.
 XX PD 07-FEB-2001.
 XX PF 28-JUL-2000; 2000EP-00116126.
 XX PR 29-JUL-1999; 99JP-00248036.
 XX PR 27-AUG-1999; 99JP-00300253.
 XX PR 11-JAN-2000; 2000JP-00118776.
 XX PR 02-MAY-2000; 2000JP-00183767.
 XX PR 09-JUN-2000; 2000JP-00241899.
 XX PA (HELI-) HELIX RES INST.
 XX PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WP1; 2001-318749/34.

PT Primer sets for synthesising polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
 XX Claim 8; SEQ ID NO 12810; 2537pp + Sequence Listing; English.
 XX The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a

CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesising polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention

XX SQ Sequence 1601 BP; 484 A; 290 C; 348 G; 479 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2 39e-211 Length: 1601
 Score: 1817.00 Matches: 350
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.72% Mismatches: 0
 Query Match: 99.78% Indels: 0
 DB: 4 Gaps: 0

US-10-698-228-1 (1-351) x AAH14924 (1-1601)

Qy 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
 Db 40 ATGGCGGACCCGGAAGCGCGGGCTGGATCAGATGAGATCATCTTCA 99
 Qy 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
 Db 100 GACACCAAGAAAGTGAATAAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 159
 Qy 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
 Db 160 CGGTTTGTTCATCTTTCCAAATCAGTACCTCATATTTGGAAATGTATATAACAGGCACAG 219
 Qy 61 AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrPheLys 80
 Db 220 GCTTCCTCTCGACAGCAGAGAGGTGCATTTATCAAGAGATCTCCCTCCTCGAACAAG 279
 Qy 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAsp 100
 Db 280 CTTAAGCAGATGAGAGTACTTTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTGAT 339
 Qy 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 Db 340 GGAATTGTAAATGAAATTTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 399
 Qy 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 Db 400 CGCTGCTTCTATGGCTTTCAAAATTTCTCATCAGAAATGTTCACTCAGAGATGACAGTTG 459
 Qy 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 Db 460 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATCAATTGAA 519
 Qy 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTyrIleAlaAspArgLys 180
 Db 520 ACCATGCCCTATTTAGAAAAAAGCAGATTTGGGCTTTGGCATGATGATAGCAGATAGAAA 579
 Qy 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluValPhePheSerGly 200
 Db 580 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 639
 Qy 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 Db 640 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCATCTTTTCC 699
 Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 Db 700 AATGAACCTCATCAGCAGAGATGAAGGACTTTCACCTGTGACTTTGCTTGTGCTGTGTTCCAA 759

QY	241	TyrLeuValAsnLysProSerGluGluArgValArgGluIleValAspAlaValLys	260
Db	760	TACTTAGTAAATAAGCGCTTCAGAAAGAAGGTCAGGAGATCATTTGTTGATGCTGTCAA	819
QY	261	IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle	280
Db	820	ATTGACGACGAGTGTTTTAAACAAGAGCTTTCGCCAGTTGGCCCTCATTTGGAGTGAATGCATT	879
QY	281	LeuMetLysGlnTyrrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer	300
Db	880	TTGATGGAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGNACTTGGATTCTCA	939
QY	301	LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys	320
Db	940	AAGTGTTCCTTCAGCAGAAATCCTTTTGATTTATGMAAACATCTCTTTAGAGGAAAA	999
QY	321	ThrAsnPhePheGluLysArgValSerGluTyrrGlnArgPheAlaValMetAlaGluThr	340
Db	1000	ACAAATTCCTTCAGAAACGAGTTTCAGAGTATCAGCGTTTTTCAGTTATGGCAGAAACC	1059
QY	341	ThrAspAsnValPheThrLeuAspAlaAspPhe	351
Db	1060	ACAGATAACGTCCTTCACCTTGGATGCAGATTTT	1092
RESULT 8			
AA	C78111	ID	AAC78111 standard; cDNA; 1989 BP.
XX	XX	AC	AAC78111;
XX	XX	DT	08-FEB-2001 (first entry)
XX	XX	DE	Human cancer associated gene sequence SEQ ID NO:505.
XX	XX	KW	Human; cancer associated gene; cancer antigen; detection; cancer;
XX	XX	KW	diagnosis; cystostatic; proliferative; vulnery; immunomodulator;
XX	XX	KW	antidiabetic; antiasthmatic; antirheumatic; antithyroid; antiviral;
XX	XX	KW	antiinflammatory; antihypertensive; antiallergic; antibacterial; cardiant;
XX	XX	KW	dermatologic; neuroprotective; thrombolytic; coagulant; nootropic;
XX	XX	KW	vasotropic; antiporiatic; antiangiogenic; gene therapy; inflammation;
XX	XX	KW	immune disorder; haematopoietic cell disorder; autoimmune disorder;
XX	XX	KW	allergic reaction; graft versus host disease; organ rejection;
XX	XX	KW	haemostatic; thrombolytic; cardiovascular disorder; infection;
XX	XX	KW	neurological disease; drug screening; ss.
XX	XX	OS	Homo sapiens.
XX	XX	PN	WO200055350-A1.
XX	XX	PD	21-SEP-2000.
XX	XX	PF	08-MAR-2000; 2000WO-US005882.
XX	XX	PR	12-MAR-1999; 99US-0124270P.
XX	XX	PA	(HUMA-) HUMAN GENOME SCI INC.
XX	XX	PI	Rosen CA, Ruben SM;
XX	XX	DR	WPI; 2000-587533/55.
XX	XX	DR	P-PSDB; AAB43902.
XX	XX	PT	Novel isolated nucleic acids comprising sequences encoding peptides
XX	XX	PT	useful for treating or diagnosing e.g. cancer.
XX	XX	PS	Claim 1; Page 1035-1036; 2352pp; English.
XX	XX	CC	AAC77607 to AAC78448 encode the human cancer associated proteins given in
XX	XX	CC	AB43398 to AB44239. The proteins can have activities based on the
XX	XX	CC	tissues and cells the genes are expressed in. Example of activities
XX	XX	CC	include: cystostatic; proliferative; vulnery; immunomodulator;
XX	XX	CC	antidiabetic; antiasthmatic; antirheumatic; antithyroid; antineoplastic;

QY 218 rPheSerAsnGluLeuLeuSerArgPheGluGluGluHisCysAspPheAlaCysLeuMe 238
 DB 845 ATTTCTTAATGAACCTTATTACGAGATGAGGGTTTACACTGTGATTTGCTTCCCTGAT 904
 QY 238 tPheGlnTyrLeuValAsnLysProSerGluGluGluArgValArgGluLeuLeuValAspAl 258
 DB 905 GTTCMAACACTGGTACACAAACCATCGGAGGAGAGTAAAGAAATAATTATCAATGC 964
 QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuLeuGlyMetAs 278
 DB 965 TGTTCGGATAGAACAGGAGTTCCTCACTGAGGCCTTCCTGTGTGAAGCTCATTTGGATGAA 1024
 QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuG 298
 DB 1025 TTGCACCTTAATGAAGCAATACATTGAGTTTGTGGCAGACAGACTTATGCTGGAACTGGG 1084
 QY 298 yPheSerLysValPheGlnAlaGluAenProPheAspPheMetGluAsnIleSerLeuG 318
 DB 1085 TTTTAGCAAGGTTTTCAGAGTAGAAGAACCCATTGACTTTATGGAGAAATAATTTCACCTGGA 1144
 QY 318 uGlyLysThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
 DB 1145 AGCAAGACTAATCTTTTCAGAGAGAGTAGGCGAGTATCAGAGGATGGGAGTGATGTC 1204
 QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
 DB 1205 AAGTCCAAACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1244

RESULT 9

ADK70302

ID ADK70302 standard; cDNA; 2215 BP.

XX AC

XX ADK70302;

XX DT

XX 06-MAY-2004 (first entry)

XX DE

XX Respiratory disease differentially expressed cDNA #38.

XX KW

XX ds; gene; cytostatic; respiratory; antiasthmatic; gene therapy;

XX differential gene expression; respiratory disorder; lung cancer;

XX chronic obstructive pulmonary disease; emphysema; asthma.

XX OS

XX Homo sapiens.

XX PN

XX W02003101283-A2.

XX XX

XX 11-DEC-2003.

XX XX

XX 02-JUN-2003; 2003WO-US017409.

XX XX

XX 04-JUN-2002; 2002US-0386005P.

XX XX

XX (INCY-) INCYTE CORP.

XX XX

XX Rickert PK, Krasnow R;

XX XX

XX WPI; 2004-042945/04.

XX XX

PT New combination comprising cDNAs and proteins that are differentially
 PT expressed in respiratory disorders, useful for diagnosing or treating
 PT respiratory diseases e.g. lung cancer, chronic obstructive pulmonary
 PT diseases or asthma.

XX XX

XX Claim 1; SEQ ID NO 38; 343pp; English.

XX CC

CC The invention relates to cDNA sequences that are differentially expressed
 CC in respiratory disorders or their complements or encoded proteins. The
 CC cDNAs and proteins are useful for diagnosing, treating or monitoring
 CC treatment of a subject with a respiratory disease including lung cancer,
 CC chronic obstructive pulmonary diseases, emphysema or asthma. The protein
 CC is also useful for screening molecules or compounds to identify at least
 CC one ligand which specifically binds the protein. It is also useful for
 CC preparing and purifying a polyclonal or monoclonal antibody. This

CC sequence corresponds to a cDNA of the invention.
 XX Sequence 2215 BP; 593 A; 476 C; 523 G; 623 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1,42e-167 Length: 2215
 Score: 1460.50 Matches: 277
 Percent Similarity: 88.42% Conservative: 36
 Best Local Similarity: 78.25% Mismatches: 35
 Query Match: 80.20% Indels: 6
 DB: 12 Gaps: 1

US-10-698-228-1 (1-351) x ADK70302 (1-2215)

QY 4 ProGluArgProGluAlaAlaGlyLeuAspGlnArgSerSerSerAspThrAs 23
 DB 188 CCTGAGGGGACCGCGCTCTGGCCAGCAAGACCGGAGGAGGATCTTCAGGAGCCAC 247
 QY 23 nGluSerGluLeuLysSer-----AsnGluGluProLeuLeuArgLysSe 38
 DB 248 GGAGCCGAAAACTAAAGCAGCTGCCCGCGGTGGAGGATGAGCCGCTCTGAGAGAAA 307
 QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysG 58
 DB 308 CCCCCTGGCTTTGTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGATAGAA 367
 QY 58 nAlaGlnAlaSerPheThrThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
 DB 368 GGCAGAGGCTTCCTTTTGGACCGCGGAGGAGGTGGACCTCTCCAGGACATTCCAGCAGT 427
 QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAl 98
 DB 428 GGAATCCCTGAAACCCGAGGAGAGATATTTATATATCCCATGTTCTTGGCTTCTTTTCAGC 487
 QY 98 aSerAspGlyIleValAsnGluValGluArgPheSerGlnGluValGlnValPr 118
 DB 488 AAGCGATGGCATAGTAAATGAAACTTGTGTGAGCGATTAGCCAGAAAGATTCCAGATTAC 547
 QY 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
 DB 548 AGAAGCCGCTGTTTCTATGCTTCCAAATGCCATGGAAACATACATCTTGAATGTA 607
 QY 138 rSerLeuLeuLeuAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
 DB 608 TAGTCTTCTTATGACACTTACATAAAGATCCCAAGAAAGAGGAATTTCTTCAATGC 667
 QY 158 aIleGluThrMetProTyrValLysLysIleAlaAspTrpAlaLeuArgTrpIleAlaAs 178
 DB 668 CATTGAAACGATGCCCTTGTGTCAAGAAAGAGCAGACTGGGCTTTCGCTGGATTGGGGA 727
 QY 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaValGluGluValPhePh 198
 DB 728 CAAAGAGGCTACCTATGTTGTAACGGTGTGTAGCCCTTGTCTGAGTGAAGAGCAATTTCTT 787
 QY 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
 DB 788 TTCGGTCTTTTGGCTCGATATTTCTGCTCAAGAAACGAGGACTGATGCTGCTGCCTCAC 847
 QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
 DB 848 ATTTTCTTAATGAACCTTATTAGCAGAGATGAGGGTTTACACTGTGATTTTGTCTGCTGAT 907
 QY 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluLeuIleValAspAl 258
 DB 908 GTTCAACACCTGGTACCAACCATCGGAGAGAGAGTAAAGAAATAATTATCAATGC 967
 QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
 DB 968 TGTTCGGATAGAACAGGAGTTCCTCACTGAGGCTTGGCTGTGAGCTCATTTGGGATGAA 1027
 QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuG 298
 DB 1028 TTGCACTCTAATGAAGCAATACATTGAGTTTGTGGCAGACAGACTTATGCTGGAAGTGGG 1087

QY 298 yPheSerLysValPheGlnAlaGluAenProPheAspPheMetGluAenLysSerLeuG1 318
Db 1088 TTTTAGCAAGGTTTTCAGAGTAGAGAACCCATTGACTTTTATGAGAGATATTTCACTGGA 1147
QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
Db 1148 AGGAAGAGCTACTCTTTGAGAGAGAGTAGGCGAGTATCAGAGAGTAGGAGTGATGTC 1207
QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1208 AAGTCCACAGAGATTTCTTTTACCTTGAGTGTGACTTC 1247
RESULT 10
ADJ56536
ID ADJ56536 standard; cDNA; 2216 BP.
XX
AC ADJ56536;
XX
DT 06-MAY-2004 (first entry)
XX
DE Murine cDNA differentially expressed in MYCN activated cells SeqID 342.
XX
KW mouse; murine; differential expression; transactivator; proto-oncogene;
KW neuroblastoma; small cell lung cancer; cytostatic; gene therapy; ss;
KW MYCN activated cell.
XX
OS Mus musculus.
XX
PN US2003119009-A1.
XX
PD 26-JUN-2003.
XX
PF 25-FEB-2002; 2002US-00084817.
XX
PR 23-FEB-2001; 2001US-0270784P.
XX
(STUA/) STUART S G.
PA (NUCH/) NUCHTERN J G.
PA (PLOW/) PLOW S E.
PA (SHOH/) SHOHET J M.
XX
PI Stuart SG, Nuchtern JG, Plon SE, Shohet JM;
DR WPI; 2003-635698/60.
XX
PT New genes regulated by MYCN activation, useful in gene therapy,
PT particularly for treating a subject with e.g. neuroblastoma or other
PT cancers, or for diagnosing, staging or monitoring the treatment of the
PT cancer.
XX
PS Claim 1; SEQ ID NO 342; 27pp; English.
XX
CC This invention relates to novel isolated cDNAs that are differentially
CC expressed in MYCN activated cells. Specifically, it refers to
CC polynucleotide sequences that exhibit differential expression patterns in
CC cells activated by the transactivator MYCN, where MYCN is a proto-
CC oncogene that is amplified in neuroblastoma cells and is common in small
CC cell lung cancers. The present invention describes these cDNA molecules
CC as useful for in hybridisation assays to detect expression of nucleic
CC acids (or complementary nucleic acids) in a present in a given sample, as
CC well as for screening assays by identifying molecules or compounds that
CC specifically bind the cDNA as a ligand and modulate function or activity.
CC Accordingly, these compositions exhibit cytostatic activity and can also
CC be used for gene therapy purposes. This polynucleotide sequence is a cDNA
CC that is differentially expressed in MYCN activated cells, given in an
CC exemplification of the invention. NOTE: This sequence does not appear in
CC the printed specification but has been obtained in electronic format from
CC the US Patent Office at
CC ftp.seqdata.uspto.gov/sequence.html?DocID=20030119009.
XX
SQ Sequence 2216 BP; 593 A; 476 C; 524 G; 623 T; 0 U; 0 Other;

Alignment Scores: 1.42e-167 Length: 2216
Pred. No.: 1460.50 Matches: 277
Score: 88.42% Conservative: 36
Percent Similarity: 88.42% Mismatches: 35
Best Local Similarity: 78.25% Indels: 6
Query Match: 80.20% Gaps: 1
DB: 10
US-10-698-228-1 (1-351) x ADJ56536 (1-2216)
QY 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
Db 189 CTTGAGCGGACCCCGCTCTGGCCAGACAGCCGAGGAGGATCTTCCAGGAGCCAC 248
QY 23 nGluSerGluLysSer-----AsnGluGluProLeuLeuArgLysSe 38
Db 249 GGAGCCGAAAACTAAAGCAGCTGCCCGCGGTGGAGGATGAGCGCTGCTGAGAGAAA 308
QY 38 rSerArgArgPheValPhePheProLleGlnTyrProAspLleTyrLysMetTyrLysG1 58
Db 309 CCCCCGCGCTTTGTCATCTTCCCATCGAGTACCATGATCTGCGCAGATGTATAGAA 368
QY 58 nAlaGlnAlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
Db 369 GGCAGAGCTTCTTTTGGACCGCGGAGAGGTGACCTCTCCAGAGGACATTCAGACTG 428
QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheLleSerHisLleLeuAlaPheAlaAl 98
Db 429 GGAATCCCTGAACCCCGAGGAGAGATATTTATATATCCATGTTCTTGGCTTTCTTTCGAGC 488
QY 98 aSerAspGlyLleValAsnGluAenLeuValGluArgPheSerGlnGluValGlnValPr 118
Db 489 AAGCGATGGCATAGTAAACTTTGGTGGAGCGATTTAGCCAAAGAAAGTTTCAGATTAC 548
QY 118 oGluAlaArgCysPheTyrGlyPheGlnLleLeuLleGluAenValHisSerGluMetTy 138
Db 549 AGAAGCCCGCTGTTTCTATGGCTTCCAAATTTGCCATAAACAATACATCTCGAAATGTA 608
QY 138 rSerLeuLeuLeuAspThrTyrLleArgAspProLysLysArgGluPheLeuPheAsnAl 158
Db 609 TAGTCTTCTTATTGACACTTACATAAAGATCCCAAGAAAGGAAATTTCTTCTCAATGC 668
QY 158 alleGluThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpLleAlaAs 178
Db 669 CATTGAAACGATGCTTGTGTCAAGAAAGAGCCAGACTGGCGCTTGCCTGGATTGGGGA 728
QY 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePh 198
Db 729 CAAAGAGGCTACCTATGGTGAACGTGTGTAGCCCTTTGCTGAGTGGAGGAGCATTTTCTT 788
QY 198 eSerGlySerPheAlaAlaLlePheTrpLeuLysLysArgGlyLeuMetProGlyLysTh 218
Db 789 TTCCGGTTCTTTTGGCTCGATATTTCTGGCTCAAGAAAGAGGACTGATGCTGGCTCAC 848
QY 218 rPheSerAsnGluLeuLysSerArgAspGluGlyLysHisCysAspPheAlaCysLeuMe 238
Db 849 ATTTTCTAATGAACCTTATTAGCAGATGAGGGTTTACACTGTGATTGCTTGCCTGAT 908
QY 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluLleLleValAspAl 258
Db 909 GTTCAAAACCTTGGTACAAACCATCGAGGAGAGAGTAAAGAGAAATAATATCAATGC 968
QY 258 aValLysLleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuLleGlyMetAs 278
Db 969 TGTTCCGATAGAACAGGAGTTCTCTAGGCGCTTGGCTGAGGCTCATTTGGATGAA 1028
QY 278 nCysLleLeuMetLysGlnTyrLleGluPheValAlaAspArgLeuLeuValGluLeuG1 298
Db 1029 TTGCACCTAATGAAGCAATACATTGATTTGTGCGCAGACAGACTTATGCTGGAACCTGG 1088
QY 298 yPheSerLysValPheGlnAlaGluAenProPheAspPheMetGluAenLleSerLeuG1 318
Db 1089 TTTTAGCAAGGTTTTCAGAGTAGAGAACCCATTGACTTTTATGAGAGATATTTCACTGGA 1148

QY 318 uGlyLysThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetal 338
 Db 1149 AGGAAGACTACTCTTTTGAAGAAGAGTAGCGAGTATCAGAGGATGGGAGTGATGTC 1208
 QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
 Db 1209 AAGTCCACAGAGATCTTTTACCTTGGATGCTGACTTC 1248

RESULT 11

AAS44917/c

ID AAS44917 standard; DNA; 2482 BP.

XX AAS44917;

AC AAS44917;

XX 18-DEC-2001 (first entry)

XX Human contig polynucleotide sequence #170.

XX Mammal; human; rhesus monkey; baker's yeast; fission yeast; Norway rat;

KW mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;

KW cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;

KW nervous system disorder; inflammatory disorder; cell differentiation; ds;

KW angiogenesis; stem cell growth factor; activin; inhibin; cartilage; burn;

KW genetic disorder; bone regeneration; tendon; ligament; tissue repair;

KW cytostatic; antirheumatic; antiarthritic; vulnary; antiinflammatory;

KW antibacterial; immunosuppressive; vasotropic; antiparkinsonian;

KW neuroprotective; osteopathic; antidiabetic; antiasthmatic; antiallergic;

KW immunostimulant; analgesic; gene therapy.

XX Homo sapiens.

OS Synthetic.

XX WO200164834-A2.

XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US004926.

XX 28-FEB-2000; 2000US-00515126.

XX 18-MAY-2000; 2000US-00577409.

XX 17-JUN-2000; 2000US-00597707.

XX 14-JUL-2000; 2000US-00616807.

XX 19-SEP-2000; 2000US-00664641.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;

XX Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;

XX Drmanac R;

XX WPI; 2001-589862/66.

XX P-PSDB; AAU28017.

XX Novel polypeptides and nucleic acids obtained from cDNA libraries

XX prepared from various human tissues for diagnosis, treatment of cancer,

XX neurological, inflammatory disorders and for use in arrays for detection.

XX Claim 1; SEQ ID NO 514; 153pp; English.

XX Sequences AAS44576-AAS44919 represent full-length polynucleotides and

XX contig polynucleotides encoding polypeptides of the invention. The DNA

XX and protein sequences are useful for the treatment, diagnosis and

XX prevention of various types of disorder in a mammalian subject such as a

XX human, dog, monkey, mouse, hamster or rat. The disorders include cancers

XX such as leukaemia, lymphoma and neuroblastoma, autoimmune disorders such

XX as multiple sclerosis, connective tissue disease, rheumatoid arthritis,

XX diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system

XX disorders such as Parkinson's disease, Alzheimer's disease, Huntington's

XX chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and

XX Wernicke disease, inflammatory disorders such as nephritis, Crohn's

XX disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory

XX bowel disease. The sequences exhibit activity relating to angiogenesis,

CC cell proliferation, cell differentiation, stem cell growth factor,
 CC activin or inhibin. Therefore, they can be used to manipulate stem cells
 CC in culture to give rise to neuroepithelial cells that can be used to
 CC augment or replace cells damaged by illness, accidental damage or genetic
 CC disorders. The sequences may also be used for regeneration of bone,
 CC cartilage, tendons and ligaments and in tissue repair and burn healing.
 CC Note: Some sequences for this patent did not form part of the printed
 CC specification, but were obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 2482 BP; 707 A; 598 C; 522 G; 655 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.68e-167 Length: 2482

Score: 1460.50 Matches: 277

Percent Similarity: 88.42% Conservative: 36

Best Local Similarity: 78.25% Mismatches: 35

Query Match: 80.20% Indels: 6

DB: 4 Gaps: 1

US-10-698-228-1 (1-351) x AAS44917 (1-2482)

QY 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23

Db 2181 CCTGAGCGGACCGCGCTCTGGCCAGCAACCGCGAGGAGGATCTCCAGGAGCCAC 2122

QY 23 nGluSerGluLeuLysSer-----AsnGluGluProLeuLeuArgLysSe 38

Db 2121 GGAGCCGAAACTAAAGCAGCTGCCCGCGGTGGAGGATGAGCGCTGTGAGAGAAA 2062

QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGl 58

Db 2061 CCCCAGCGCTTGTCTATCTCCCATGAGTACCATGATCTGGCAGATGATATAAGAA 2002

QY 58 nAlaGlnAlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78

Db 2001 GGACAGAGCTTCTTTGGACCGCGAGGAGGTGACCTCTCCAAAGGACATTCAGACTG 1942

QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAl 98

Db 1941 GGAATCCCTGAACCGGAGGAGATATTTATATATCCCATGTTCTGGCTTCTTTGCGAGC 1882

QY 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118

Db 1881 AAGCATGGCATAGTAATAAGAACTTGTGTGGAGCGATTAGCCAAAGAAAGTTTCAGATTAC 1822

QY 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuLeuGluAsnValHisSerGluMetTy 138

Db 1821 AGAAGCCCGCTGTCTATGCTTCCAAATTTGCCATGGGAAACATACATCTTCTGAATGPA 1762

QY 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158

Db 1761 TAGTCTTCTTATGACACTTACATAAAGATCCCAAGAAAGGGAATTTCTTCTCATGTC 1702

QY 158 alieGluThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178

Db 1701 CATTAACAGATGCTTGTGTCAAGAGAGGAGACTGGGCTTGGCTGGATTGGGGA 1642

QY 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaValGluGluValPhePh 198

Db 1641 CAAAGAGCTACCTATGTTGTAACCGTGTGTGTAGCTTCTGCTGAGTGAAGGATTTCTT 1582

QY 198 eSerGlySerPheAlaIlePheThrLeuLysLysArgGlyLeuMetProGlyLeuTh 218

Db 1581 TTCGGTCTTCTTTCGTCGATATTTCTGCTCAAGAAACGAGGACTGATGCTCGGCTCAC 1522

QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLysHisCysAspPheAlaCysLeuMe 238

Db 1521 ATTTTCTAATGAATCTTATAGCAGATGAGGTTTACCTGTGATTTGTTGCTGAT 1462

QY 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258

Db 1461 GTTCAACACCTGTTACACAAACCATCGGAGGAGAGAGTAAAGAGAAATAATTATCAATGC 1402

```
QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
Db 1401 TGTTCGATAGAACAGAGTTCCTCACTGAGGCTTGGCCCTGTGAAGCTCATTTGGGATGAA 1342
QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuG1 298
Db 1341 TTGCATCTTAATGAAGCAATACATTGAGTTTGTGGCAGACAGACTTATGCTGGAATGGG 1282
QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG1 318
Db 1281 TTTTAGCAAGGTTTTCAGAGTAGAGAACCAATTTGACTTTATGAGATATTTCACTGGA 1222
QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
Db 1221 AGGAAAGACTAATCTCTTTGAGAGAGAGTAGGCGAGTATCAGAGATGGAGTGATGTC 1162
QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1161 AAGTCCAAACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1122

RESULT 12
ABL65414
ID ABL65414 standard; DNA; 2500 BP.
AC ABL65414;
XX
DT 15-MAY-2002 (first entry)
XX
DE Lung cancer related gene sequence SEQ ID NO:3751.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; Gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US010838.
XX
PR 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0233133P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 26-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
```

```
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
XX (AVAL-) AVALON PHARM.
PA
XX Young PE, Augustus M, Carter KC; Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX
PS Claim 1; SEQ ID NO 3751; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC tumour
XX
SQ Sequence 2500 BP; 679 A; 522 C; 595 G; 704 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.7e-167 Length: 2500
Score: 1460.50 Matches: 277
Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
DB: Gaps: 1

US-10-698-228-1 (1-351) x ABL65414 (1-2500)
QY 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
Db 302 CCTGAGCGGAGCCCGGCTCTCTGCCAGACGCGGAGGAGATCTTCCAGGAGCCAC 361
QY 23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
Db 362 GGAGCCGAAACTAAAGCAGCTGCCCGCGCGTGGAGGATGAGCGCTGCTGAGAGAAA 421
QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysG1 58
Db 422 CCCCCCGCGCTTTGTCTATCTTCCCATGAGTACCATGATATCTGGCAGATGTTAAGAA 481
QY 58 nAlaGlnAlaSerPheThrThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
Db 482 GGCAGAGGCTTCTCTTTGGACCGCGGAGGAGGTGACCTCTCCAAAGGACATTGACACTG 541
```

Qy 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaA 98
Db 542 GGAATCCCTGAAACCCGAGGAGAGATATTTATATCCCATGCTTTCCTTTGCAGC 601
Qy 98 aSerAspGlyLysValAsnGluLysValGluArgPheSerGlnGluValGlnValPr 118
Db 602 AAGCGATGGCAGATGTAATGAAACTTGGTGGAGCATTTAGCCAAAGATTCAGATTAC 661
Qy 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
Db 662 AGAAGCCCGCTGTTCTATGCTTCCAAATTCGCATGCGAACAATACATCTCTGAATGTA 721
Qy 138 rSerLeuLeuLeuAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
Db 722 TAGTCTTCTTATTGACACTTACATAAAAGATCCCAAGAAAGGAATTTCTCTTCAATGC 781
Qy 158 aIleGluThrMetProTyrValLysLysLysAlaAspTyrAlaLeuArgTyrIleAlaAs 178
Db 782 CATTGAAACGATGCTTGTGTCAAGAAAGAGCGAGACTGGCGCTTGGCTGGATTTGGGGA 841
Qy 178 pArgLysSerThrPheGlyGluArgValAlaAlaPheAlaValGluGlyValPhePh 198
Db 842 CAAGAGCGTACCTATGCTGACGTGTGTAGCCCTTGTGTCAGTGAAGGCATTTCTT 901
Qy 198 eSerGlySerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuTh 218
Db 902 TTCGGTTCCTTTTGGCGTCGATATTTCTGGCTCAAGAAAGAGGACTGATGCTGCCTCAC 961
Qy 218 rPheSerAsnGluLeuLysSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
Db 962 ATTTCTTAATCAACTTATTAGCAGATGAGGGTTTACACTGTGATTTTGTCTTGCCTGAT 1021
Qy 238 tPheGluTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
Db 1022 GTTCAAAACCTGTGACACAAACCATCGAGGAGAGATAGAGAAATATTATCAATGC 1081
Qy 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
Db 1082 TGTTCCGATAGACAGAGATTCCTCACTGAGGCCCTTGCCTGTGAAGCTCATTTGGATGAA 1141
Qy 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGl 298
Db 1142 TTGCACTTAATGAAGCAATACATTGATTTGTGGCAGACAGACTTATGCTGGAACCTGG 1201
Qy 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGl 318
Db 1202 TTTTAGCAAGGTTTTCAGAGTAGAGAACCCATTTGACTTTATGAGAGATATTTCACTGGA 1261
Qy 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
Db 1262 AGGAAAGACTTAATCTTTTGAGAGAGAGTAGGCGAGTATCAGAGGATGGAGTGTGTC 1321
Qy 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1322 AAGTCCAAACAGAGATTTCTTTTACCTTTGGATGCTGACTTC 1361

RESULT 13

ABL66517

ID ABL66517 standard; DNA; 2500 BP.

AC ABL66517;

XX ABL66517;

XX 15-MAY-2002 (first entry)

DT Lung cancer related gene sequence SEQ ID NO:4854.

DE Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytosclastic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.

OS Homo sapiens.

XX

PN WO200194629-A2.
XX 13-DEC-2001.
XX 30-MAY-2001; 2001WO-US010838.
XX 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0233133P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 26-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.

(AVAL-) AVALON PHARM.

Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
Soppet DR, Weaver Z;

WPI; 2002-188264/24.

Screening for anti-neoplastic agent involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, and determining a change in expression of a gene of a signature gene set.

Claim 1; SEQ ID NO 4854; 44pp; English.

The present invention describes a method (M1) for screening for an anti-neoplastic agent. The method involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, determining a change in expression of at least one gene (I) of a signature gene set, where (I) comprises a sequence (S) selected from 8447 sequences (given in ABL61664 to ABL70110), or is at least 95% identical to (S), where a change in expression is indicative of anti-neoplastic activity. (I) has cytostatic activity and can be used in gene therapy. M1 can be used for screening an anti-neoplastic agent, and can be used for producing a product which is the data collected with respect to the anti-neoplastic agent as a result of M1, and the data is sufficient to convey the chemical structure and/or

CC properties of the agent. M1 can be used in the treatment of cancer such
 CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
 CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
 CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
 CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
 CC tumour
 XX
 SQ Sequence 2500 BP; 679 A; 522 C; 595 G; 704 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1.7e-167 Length: 2500
 Score: 1460.50 Matches: 277
 Percent Similarity: 88.42% Conservative: 36
 Best Local Similarity: 78.25% Mismatches: 6
 Query Match: 80.20% Indels: 6
 DB: Gaps: 1

US-10-698-228-1 (1-351) x ABL66517 (1-2500)

QY 4 ProGluArgProGluAlaAlaGlyLeuAspGluAspGlu-ArgSerSerSerThrAs 23
 DB 302 CCTGAGCGGACCCGGGCTCTGCGCCAGCAAGACCGGAGGAGGATCTTCCAGGAGCCAC 361
 QY 23 nGluSerGluLeuLeuSer-----AsnGluGluProLeuLeuArgLysSe 38
 DB 362 GGAGCGGAAACTAAAGCAGCTCCCGCGCGTGGAGGATGAGCGCTGCTGAGAGAAA 421
 QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTyrIlePheGly 58
 DB 422 CCCCAGCGCTGTCATCTCCCATCGAGTACCATGATATCTGCGAGATGATAAGAA 481
 QY 58 nLaGlnAlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
 DB 482 GGCAGAGGCTTCTTTTGGACCGCGGAGGAGTGTGACCTCTCCAGGACATTCAGCACTG 541
 QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAla 98
 DB 542 GGAATCCCTGAAACCGGAGGAGAGATATTTATATCCATGTTCTGCGCTTCTTTCGAGC 601
 QY 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
 DB 602 AGCGATGCGATAGTAAGTAAGAAACTTGTGGAGCGATTTAGCCAGAGTTCAGATTAC 661
 QY 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
 DB 662 AGAAGCGCGCTGTTCTATGCTTCCAAATTGCCATGCAAAACATACATTCGAAATGTA 721
 QY 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
 DB 722 TAGTCTTCTTATTGACACTTACATAAAGATCCCAAGAAAGGGAATTTCTCTTCAATGC 781
 QY 158 aIleGluThrMetProTyrValLysLysLysAlaAspTyrAlaLeuArgTyrIleAlaAs 178
 DB 782 CATTAAGACATGCTTGTGTCAAGAGAGAGGAGAGCTGGCGCTTGGCTGGATTGGGA 841
 QY 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePh 198
 DB 842 CAAAGAGGCTACCTATGTTGAACTGTTGTAGCTTTGCTGCTGAGTGGAGGCAATTTCTT 901
 QY 198 eSerGlySerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuTh 218
 DB 902 TTCCGGTCTCTTTCGCTGATATTTGCTTCAAGAAACGAGGACTGATGCTGCGCTCAC 961
 QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
 DB 962 ATTTCTAATGAATCTATTAGCAGATGAGGTTTACATGATTTTCTGCTGCTGAT 1021
 QY 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
 DB 1022 GTTCAACACCTGTTACCAACCATCGGAGGAGAGTAAGAGAAATATATCAATGC 1081
 QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278

DB 1082 TGTTCGATAGAACAGAGTTCTCTCACTGAGGCCTTGCCTGTGAAGCTCATTTGGATGAA 1141
 QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGl 298
 DB 1142 TTGCACCTTAATGAAGCAATACATTTAGTTTGTGCGACAGACACTTATGCTGGAACGG 1201
 QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGl 318
 DB 1202 TTTTAGCAAGGTTTTCAGAGTAGAGAACCCCAATTTGACTTTTATGGAGAAATATTTCACTGA 1261
 QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGlnTyrGlnArgPheAlaValMetAl 338
 DB 1262 AGGAAAGACTAATCTTTTGAAGAGAGTAGGCGAGTATCAGAGGATGGGAGTGTATGTC 1321
 QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
 DB 1322 AAGTCCAAACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1361

RESULT 14
 ABL65859
 ID ABL65859 standard; DNA; 2500 BP.
 XX
 AC ABL65859;
 XX
 DT 15-MAY-2002 (first entry)
 XX
 DE Lung cancer related gene sequence SEQ ID NO:4196.
 XX
 KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytosstatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200194629-A2.
 XX
 PD 13-DEC-2001.
 XX
 PF 30-MAY-2001; 2001WO-US010838.
 XX
 PR 05-JUN-2000; 2000US-0209473P.
 PR 05-JUN-2000; 2000US-0209531P.
 PR 18-SEP-2000; 2000US-0233133P.
 PR 18-SEP-2000; 2000US-0233617P.
 PR 20-SEP-2000; 2000US-0234009P.
 PR 20-SEP-2000; 2000US-0234034P.
 PR 20-SEP-2000; 2000US-0234052P.
 PR 22-SEP-2000; 2000US-0234509P.
 PR 22-SEP-2000; 2000US-0234567P.
 PR 25-SEP-2000; 2000US-0234923P.
 PR 25-SEP-2000; 2000US-0234924P.
 PR 25-SEP-2000; 2000US-0235077P.
 PR 25-SEP-2000; 2000US-0235082P.
 PR 25-SEP-2000; 2000US-0235134P.
 PR 25-SEP-2000; 2000US-0235280P.
 PR 26-SEP-2000; 2000US-0235637P.
 PR 26-SEP-2000; 2000US-0235638P.
 PR 27-SEP-2000; 2000US-0235711P.
 PR 27-SEP-2000; 2000US-0235720P.
 PR 27-SEP-2000; 2000US-0235840P.
 PR 27-SEP-2000; 2000US-0235863P.
 PR 28-SEP-2000; 2000US-0236028P.
 PR 28-SEP-2000; 2000US-0236032P.
 PR 28-SEP-2000; 2000US-0236033P.
 PR 28-SEP-2000; 2000US-0236034P.
 PR 28-SEP-2000; 2000US-0236109P.
 PR 28-SEP-2000; 2000US-0236111P.
 PR 29-SEP-2000; 2000US-0236842P.
 PR 29-SEP-2000; 2000US-0236891P.
 PR 02-OCT-2000; 2000US-0237172P.
 PR 02-OCT-2000; 2000US-0237173P.
 PR 02-OCT-2000; 2000US-0237278P.

PR 02-OCT-2000; 2000US-0237294P.
 PR 02-OCT-2000; 2000US-0237295P.
 PR 02-OCT-2000; 2000US-0237316P.
 PR 03-OCT-2000; 2000US-0237425P.
 PR 03-OCT-2000; 2000US-0237598P.
 PR 03-OCT-2000; 2000US-0237604P.
 PR 03-OCT-2000; 2000US-0237606P.
 PR 03-OCT-2000; 2000US-0237608P.
 PR 01-NOV-2000; 2000US-0244867P.
 PR 01-NOV-2000; 2000US-0245084P.
 XX (AVAL-) AVALON PHARM.
 XX
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 XX WPI; 2002-188264/24.
 XX
 PT Screening for anti-neoplastic agent involves exposing cells to a chemical
 PT agent to be tested for anti-neoplastic activity, and determining a change
 PT in expression of a gene of a signature gene set.
 XX
 PS Claim 1; SEQ ID NO 4196; 44pp; English.
 XX
 CC The present invention describes a method (M1) for screening for an anti-
 CC neoplastic agent. The method involves exposing cells to a chemical agent
 CC to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening an
 CC anti-neoplastic agent, and can be used for producing a product which is
 CC the data collected with respect to the anti-neoplastic agent as a result
 CC of M1, and the data is sufficient to convey the chemical structure and/or
 CC properties of the agent. M1 can be used in the treatment of cancer such
 CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
 CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
 CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
 CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
 CC tumour
 XX
 SQ Sequence 2500 BP; 679 A; 522 C; 595 G; 704 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.7e-167 Length: 2500
 Score: 1460.50 Matches: 277
 Percent Similarity: 88.42% Conservative: 36
 Best Local Similarity: 78.25% Mismatches: 35
 Query Match: 80.20% Indels: 6
 DB: 6 Gaps: 1
 US-10-698-228-1 (1-351) x ABL65859 (1-2500)
 QY 4 ProGluArgProGluAlaAlaGlyLeuAspGluArgSerSerSerAspThrAs 23
 DB 302 CTGTAGCGGGACCGCGTCTGGCGAGCAAGACCGCGAGAGGATCTTCAGAGGCCAC 361
 QY 23 nGluSerGluIleLeuSer-----AsnGluGluProLeuLeuArgLysSe 38
 DB 362 GGAGCGGAAACTAAGCAGCTGCCCGCGGTGGAGGATGAGCCGCTGCTGAGAGAAA 421
 QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTyrPheMetTyrLysG 58
 DB 422 CCCCAGCGCGTCTGTGTCATCTCCCATCGAGTACCATGATATCTGCAGATGTATAGAA 491
 QY 58 nAlaGlnAlaSerPheTyrThrAlaGluValAspLeuSerLysAspLeuProHisTr 78
 DB 482 GCAGAGGGTCTCTTTTGACCGCGCGAGGAGGTGACCTCTCCAGAGCAATTCAGCACTG 541
 QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAla 98
 DB 542 GGAATCCCTGAAACCGGAGGAGATATTTTATATCCATGTTCTGCTTCTTTCAGC 601

QY 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
 DB 602 AAGCGATGGCATAGTAATAAAGAACTTGGTGGAGCGATTTAGCCAAAGATTCAGATTAC 661
 QY 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
 DB 662 AGAAGCCCGCTGTTTCTATGGCTTCCAAATTTGCCAATGGAACATACATCTTCTCAATGC 721
 QY 138 rSerLeuLeuLeuAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
 DB 722 TAGTCTTCTTATGACACTTACATAAAGATCCCAAGAAAGGGAATTTCTTCTCAATGC 781
 QY 158 aileGluThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
 DB 782 CATTGAAACGATGCTTGTGTCAAGAAAGCGAGACTGGCGCTTGGCTGGATTGGGA 841
 QY 178 pArgLysSerThrPheGlyGluArgValAlaPheAlaValGluGlyValPhePh 198
 DB 842 CAAAGAGGCTACCTATGCTGAACGTGTGTAGCCTTTGCTCAGTGGAGGATTTTCTT 901
 QY 198 eSerGlySerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuTh 218
 DB 902 TTCCGGTCTTTTGGCTGATATTTCTGGCTCAAGAAACGAGACTGATGCTGGCTCAC 961
 QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
 DB 962 ATTTTCTAATGAACCTATTAGCAGAGATGAGGGTTTACACTGTGATTTTGTGCTGAT 1021
 QY 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
 DB 1022 GTTCAACACCTGTCACAAACCATCGAGAGAGAGTAAGAGAAATAATATCAATGC 1081
 QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
 DB 1082 TGTTCGATGAACAGAGGTTCTCTACCTGAGGCTTGTGCTGAAAGCTCATTTGGGATGA 1141
 QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuG 298
 DB 1142 TTGCACCTTAATGAAGCAATACATTTAGTTGTGGCAGACAGACTTATGCTGGAATGG 1201
 QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG 318
 DB 1202 TTTTAGCAAGGTTTTCAGAGTAGAGAACCCATTTGACTTTATGGAGATATTTCACTGA 1261
 QY 318 uGlyLysThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
 DB 1262 AGGAAAGACTAACTTCTTTGAGAAAGAGAGTAGGCGGAGTATCAGAGGATGGGAGTGC 1321
 QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
 DB 1322 AAGTCCAACAGAGANTTTCTTTACCTTGGATGCTGACTTC 1361
 RESULT 15
 ABX10335
 ID ABX10335 standard; DNA; 2500 BP.
 XX
 AC ABX10335;
 XX
 DT 28-JAN-2003 (first entry)
 XX
 DE DNA encoding protein differentially regulated in prostate cancer #4.
 XX
 KW Prostate cancer; gene expression; differential regulation;
 KW molecular marker; drug target; cancer detection; cancer diagnosis;
 KW cancer staging; cancer grading; cancer assessing; cancer monitoring;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200281638-A2.
 XX
 PD 17-OCT-2002.

XX 08-APR-2002; 2002WO-US010824.
 XX 06-APR-2001; 2001US-0281731P.
 PR 06-APR-2001; 2001US-0281732P.
 XX (ORIG-) ORIGENE TECHNOLOGIES INC.
 PA
 XX
 PI Sun Z, Jay G;
 XX WPI: 2003-058520/05.
 DR P-PSDB; ABU07433.
 DR
 XX
 PT Novel genes which are differentially regulated in prostate cancer, useful
 PT for diagnosing prostate cancer in prostate tissue sample and assessing
 PT therapeutic or preventive intervention in prostate cancer patients.
 XX
 PS Claim 1; Page 71-72; 416pp; English.
 XX
 CC The invention describes genes (I) which are differentially regulated in
 CC prostate cancer. (I) Is useful for diagnosing a prostate cancer in a
 CC sample comprising prostate tissue, which involves determining the number
 CC of target genes which are differentially-regulated in the sample, where
 CC the number is indicative of the probability that the sample comprises
 CC prostate cancer. (I) Is useful for assessing a therapeutic or preventive
 CC intervention in a subject having a prostate cancer, which involves
 CC determining the expression levels in a sample comprising prostate tissue
 CC of target genes which are differentially-regulated in prostate cancer.
 CC Preferably, the expression levels of at least 10 genes are determined.
 CC (I) is also useful for identifying agents that modulate a biological
 CC activity of a polypeptide differentially-regulated in prostate cancer
 CC cells, which involves contacting a polypeptide differentially-regulated
 CC in prostate cancer cells with a test agent under conditions effective for
 CC the test agent to modulate a biological activity of the polypeptide, and
 CC determining whether the test agent modulates the biological activity. (I)
 CC is useful as molecular markers, as drug targets, and for detecting,
 CC diagnosing, staging, grading, assessing, monitoring, prognosticating,
 CC preventing or treating, determining predisposition to diseases and
 CC conditions especially relating to prostate cancer. (I) and its expression
 CC products are used in the diagnostic test to assay for presence of cancer
 CC e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in
 CC blood etc. (I) is useful for assessing cancer e.g., to determine the type
 CC of cancer, its stage of development, the nature of genetic defect, etc.
 CC The polypeptide encoded by (I) can be used as target for therapy or drug
 CC discovery. (I) can also be used for expressing the polypeptide and thus
 CC for searching specific binding partners of the polypeptide. (I) is useful
 CC in therapeutic applications to treat prostate cancer. The identification
 CC of specific genes, and groups of genes, expressed in pathways
 CC physiologically relevant to prostate cancer permits the definition of
 CC functional and disease pathways and the delineation of targets in these
 CC pathways which are useful in diagnostic, therapeutic, and clinical
 CC applications. This sequence encodes a protein differentially regulated in
 CC prostate cancer
 XX
 SQ Sequence 2500 BP; 679 A; 522 C; 595 G; 704 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1,7e-167 Length: 2500
 Score: 1460.50 Matches: 277
 Percent Similarity: 88.42% Conservative: 36
 Best Local Similarity: 78.25% Mismatches: 35
 Query Match: 80.20% Indels: 6
 DB: 8 Gaps: 1
 US-10-698-228-1 (1-351) x ABX10335 (1-2500)
 QY 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSer-Ser-SerAspThrAs 23
 Db 302 CTGACGGGGACCGCGTCTCTGCGCAGCAAGACCGCGGAGGAGTCTTCCAGGAGCCAC 361
 QY 23 nGluSerGluileLysSer-----AsnGluGluProLeuLeuArgLysSe 38
 Db 362 GGAGCCGAAACTAAAGCAGCTGCCCCCGCGTGGAGGATGAGCGCGTCTGTGAGAGAA 421

QY 38 rSerArgArgPheValIlePhePheProIleGlnTyProAspIleTrpLysMetTyLysGl 58
 Db 422 CCCCCCGCGCTTGTGTCATCTTCCCATCGAGTATCATGATCTCGCAGATGATATAAGAA 481
 QY 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
 Db 482 GCGAGAGGCTTCTCTTTTGGACCGCGGAGGATGACCTCTCCAAAGGACATTCAGCATG 541
 QY 78 pAsnLysLeuLysAlaAspGluLysTyPheIleSerHisIleLeuAlaPhePheAlaAl 98
 Db 542 GGRATCCCTGAAACCCGAGGAGATATTTATATCCATGTTCTGGCTTCTTTTCGAGC 601
 QY 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
 Db 602 AAGCGATGGCATAGTAATAAAGCTTGGTGAGCGATTTAGCCCAAGAAAGTTTCAGATTAC 661
 QY 118 oGluAlaArgCysPheTyrglyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
 Db 662 AGAAGCCCGCTGTTTCTATGCTTCCAAATTTGCCATGGAAACATACATTTCTGAATGTA 721
 QY 138 rSerLeuLeuIleAspThrTyrlleArgAspProLysLysArgGluPheLeuPheAsnAl 158
 Db 722 TAGTCTTCTTATGACACTTACATAAAGATCCCAAGAAAGGAAATTTCTTTCATGTC 781
 QY 158 alieGluThrMetProTyrrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
 Db 782 CATTGAAACGATGCTTGTGTCAAGAAAGAGGAGACTGGCGCTTGGCTGGATGGGA 841
 QY 178 pArgLysSerThrPheGlyGluArgValAlaPheAlaAlaValGluGluValPhePh 198
 Db 842 CAAGAGGCTACCTTGTGTAACGCTTGTAGCTTGTCTCAGTGGAGGATTTCTTCTT 901
 QY 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
 Db 902 TTGCGHTTCTTTCGTCGATTTCTGCTCAAGAAACGAGACTGATGCTGGCTCAC 961
 QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
 Db 962 ATTTTCTAATGAACCTTATAGCAGAGATGAGGGTTTACACTGTGATTTTGTGCTGAT 1021
 QY 238 tPheGlnTyrlleuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
 Db 1022 GTTCAAAACACTGTGTACACAAACCATCGGAGGAGAGTAAAGAGAAATAATATCAATGC 1081
 QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
 Db 1082 TGTTGGATAGAACAGGAGTTCCTCATCTGAGGCTTGGCTGTGAAGCTCATTTGGATGAA 1141
 QY 278 nCysIleLeuMetLysGlnTyrlleGluPheValAlaAspArgLeuLeuValGluLeuGl 298
 Db 1142 TTGCACCTCTAATGAAGCAATACATTGAGTTTGTGGCAGACAGACTTATGCTGGAACGG 1201
 QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGl 318
 Db 1202 TTTTAGCAAGGTTTTCAGAGTAGAGAAACCCATTTGACTTTATGGAGAATAATTTCACTGA 1261
 QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrglnArgPheAlaValMetAl 338
 Db 1262 AGGAAGACTAATCTTCTTGGAAAGAGAGTATGAGGAGTATCAGAGGATGGGAGTATGTC 1321
 QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
 Db 1322 AAGTCCAAACAGAGAAATTTCTTTTACCTTGGATGCTGACTTC 1361

Search completed: June 16, 2005, 11:40:40
 Job time : 660 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 16, 2005, 09:43:32 ; Search time 5171 Seconds

(without alignments)
3289.070 Million cell updates/sec.

Title: US-10-698-228-1

Perfect score: 1821

Sequence: 1 MGDPERPEAGLDQDERSS.....QRFVMAETTDNVFTLDADF 351

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DRV=xlh
-Q=/cgn2_1/USPTO.spool/US10698228/runat_14062005_161416_28295/app_query.fasta_1.519
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10698228 @cgn_1_4200 @runat_14062005_161416_28295 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1821	100.0	1053	6	AR454868
2	1821	100.0	1053	6	AR454877
3	1821	100.0	1053	6	BD064764
4	1821	100.0	1053	6	BD064773

5	1821	100.0	1053	6	BD093077	BD093077	Novel pro
6	1821	100.0	1053	6	BD093086	BD093086	Novel pro
7	1821	100.0	1056	6	CQ714252	CQ714252	Sequence
8	1821	100.0	1056	9	AB166671	AB166671	Homo sapi
9	1821	100.0	1081	6	AR454870	AR454870	Sequence
10	1821	100.0	1081	6	BD064766	BD064766	Novel pro
11	1821	100.0	1081	6	BD093079	BD093079	Novel pro
12	1821	100.0	4955	6	AR454869	AR454869	Sequence
13	1821	100.0	4955	6	BD064765	BD064765	Novel pro
14	1821	100.0	4955	6	BD093078	BD093078	Novel pro
15	1821	100.0	4955	9	AB036063	AB036063	Homo sapi
16	1817	99.8	1601	6	BD156916	BD156916	Primer fo
17	1817	99.8	1601	6	AX877905	AX877905	Sequence
18	1817	99.8	1601	9	AK001965	AK001965	Homo sapi
19	1725	94.7	4532	10	BC058103	BC058103	Mus muscu
20	1627.5	89.4	171737	10	AC122379	AC122379	Mus muscu
21	1514	83.1	900	9	AB163437	AB163437	Homo sapi
22	1474	80.9	4571	9	BC042468	BC042468	Homo sapi
23	1460.5	80.2	1649	9	BC001886	BC001886	Homo sapi
24	1460.5	80.2	1653	9	BC030154	BC030154	Homo sapi
25	1460.5	80.2	1794	9	S40301	S40301	Homo sapien
26	1460.5	80.2	2500	6	AR300103	AR300103	Sequence
27	1460.5	80.2	2500	6	AR380825	AR380825	Sequence
28	1460.5	80.2	2500	6	AR38503	AR38503	Sequence
29	1460.5	80.2	2500	6	AR562011	AR562011	Sequence
30	1460.5	80.2	2500	6	AX333242	AX333242	Sequence
31	1460.5	80.2	2500	6	AX333687	AX333687	Sequence
32	1460.5	80.2	2500	6	AX334345	AX334345	Sequence
33	1460.5	80.2	2500	9	HSRR255	HSRR255	Sequence
34	1456.5	80.0	3039	9	AK092671	AK092671	Homo sapi
35	1445.5	79.4	3203	9	AK123010	AK123010	Homo sapi
36	1444	79.3	1538	5	CR762018	CR762018	Xenopus t
37	1444	79.3	1611	5	CR848103	CR848103	Xenopus t
38	1443	79.2	2111	10	MUSNRM2A	M14223	Mouse ribon
39	1439	79.0	1406	5	CR353642	CR353642	Gallus ga
40	1439	79.0	1647	5	BC080161	BC080161	Xenopus t
41	1438	79.0	217774	2	AC120603	AC120603	Rattus no
42	1431	78.6	1605	5	BC047975	BC047975	Xenopus 1
43	1429	78.5	1170	6	E14825	E14825	cdna encodi
44	1428.5	78.4	1600	5	BC041209	BC041209	Xenopus 1
45	1425.5	78.3	1328	5	DRU57965	U57965	Danio rerio

ALIGNMENTS

RESULT 1	AR454868	Sequence 2 from patent US 6682917.	1053 bp	DNA	linear	PAT 20-FEB-2004
LOCUS	AR454868	Sequence 2 from patent US 6682917.				
DEFINITION	AR454868	Sequence 2 from patent US 6682917.				
ACCESSION	AR454868	Sequence 2 from patent US 6682917.				
VERSION	AR454868.1	GI:42688823				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	1 (bases 1 to 1053)					
AUTHORS	Nakamura,Y., Arakawa,H. and Tanaka,H.					
TITLE	Protein having a ribonucleotide Reductase activity and a DNA thereof					
JOURNAL	Patent: US 6682917-A 2 27-JAN-2004;					
FEATURES	Location/Qualifiers					
source	1..1053					
	/organism="unknown"					
	/mol_type="genomic DNA"					

ORIGIN

Alignment Scores:	1.01e-170	Length:	1053
Pred. No.:	1821.00	Matches:	351
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	100.00%	Gaps:	0
DB:	6		

US-10-698-228-1 (1-351) x AR454868 (1-1053)

Qy	1	MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer	20
Db	1	ATGGCGACCCGAAAGAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA	60
Qy	21	AspThrAsnGluSerGluLeuLysSerAsnGluProLeuLeuArgLysSerSerArg	40
Db	61	GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTCTCGC	120
Qy	41	ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln	60
Db	121	CGGTTTGTGTCATCTTCCAAATCCAGTACCCCTGATATTTGGAAATGTATAAACAGCGACAG	180
Qy	61	AlaSerPheTrpThrAlaGluGluValAlaAspLysSerLysAspLeuProHisTrpAsnLys	80
Db	181	GCTTCTCTTCGGACAGCAGAAGAGGTGCACTTATCAAGGATCTCCCTCACTGGAACAAG	240
Qy	81	LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp	100
Db	241	CTTAAAGCAGATGAGAAAGTACTTTCATCTCTCACATCTTAGCCCTTTTTCGAGCCAGTGCAT	300
Qy	101	GlyIleValAsnGluAenLeuValGluArgPheSerGlnGluValGlnValProGluAla	120
Db	301	GGAAATTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGCT	360
Qy	121	ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu	140
Db	361	CGCTGTTCTATGGCTTTCAAATTCATCGAGATGCTTCACTCAGAGATGTACAGTTTG	420
Qy	141	LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu	160
Db	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA	480
Qy	161	ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys	180
Db	481	ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCCCTTGGCGATGGATAGCAGATAGAAAA	540
Qy	181	SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly	200
Db	541	TCTACTTTTGGGAAAGAGTGGTGGCCCTTGTCTGTGTAGAGAGATTTCTTCTCAGGA	600
Qy	201	SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer	220
Db	601	TCCTTTTGTGCTATATTTCTGGCTAAAGAAAGAGAGGTCTTATGCCAGGACTCACATTTTTC	660
Qy	221	AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln	240
Db	661	AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTGTCTGCCCTGATGTCCAA	720
Qy	241	TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys	260
Db	721	TACTTAGTAATAATAGCCCTTCAGAAAGAAGGTTCAGGAGATCATTTGTGTGCTGTCAAA	780
Qy	261	IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle	280
Db	781	ATTGAGCAGGAGTTTTTAAACAGAAGCCTTGCCAGTTGSCCTCATTTGGAATGATTTGCATT	840
Qy	281	LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuGlyPheSer	300
Db	841	TTGATGAAACAGTACATTTGATTTGTAGCTGCACAGATTACTTGTGTGGAACCTTGGATTCTCA	900
Qy	301	LysValPheGlnAlaGluAenProPheAspPheMetGluAsnIleSerLeuGluGlyLys	320
Db	901	AAGGTTTTTTCAGGCAGAAAAATCTTTTTGATTTTATGGAATAACATTTCTTTAGAAAGGAAA	960
Qy	321	ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr	340
Db	961	ACAAATTTCTTTCAGAAACGAGTTTCAGATATCAGCGTTTTTCAGTATGCGCAGAAACC	1020
Qy	341	ThrAspAsnValPheThrLeuAspAlaAspPhe	351

Db 1021 ACAGATAACGTCCTTCACCTTGGATGCAGATTTT 1053

RESULT 2
 AR454877 1053 bp DNA linear PAT 20-FEB-2004
 LOCUS
 DEFINITION Sequence 12 from patent US 6682917.
 ACCESSION AR454877
 VERSION AR454877.1 GI:42688832
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE
 AUTHORS 1 (bases 1 to 1053)
 TITLE Nakamura,Y., Arakawa,H. and Tanaka,H.
 Protein having a ribonucleotide Reductase activity and a DNA
 thereof
 JOURNAL Patent: US 6682917-A 12 27-JAN-2004;
 FEATURES Location/Qualifiers
 source 1..1053
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN

Alignment Scores:

Pred. No.:	1.01e-170	Length:	1053
Score:	1821.00	Matches:	351
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-698-228-1 (1-351) x AR454877 (1-1053)

Qy	1	MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer	20
Db	1	ATGGCGCACCGGAAGCCGGAGCGCGGGCTGGATCAGGATGAGATCATCTTCA	60
Qy	21	AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerSerArg	40
Db	61	GACACCAACGAAAGTGAAATGAAGTCAAAATGAAGAGCCACTCTCTTAAGAAAGAGTTC	120
Qy	41	ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln	60
Db	121	CGGTTTGTCTATCTTCCATCCAGTACCTGTATATTGGAAATGTATAAACAGCCACAG	180
Qy	61	AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys	80
Db	181	GCCTTCCTTCGGACAGCAGAGAAGAGTTGACTTATCAAGAGGATCTCCCTCACTGGAACAAG	240
Qy	81	LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp	100
Db	241	CTTAAAGCAGATGAAGATGACTTCACTCTCACATCTTAGCCCTTTTTCAGCCAGCGTAT	300
Qy	101	GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla	120
Db	301	GGAAATCTGAATGAAATTTGGTGGACGCTTTAGTCAGAGGTGCAGGTTCACAGAGCT	360
Qy	121	ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerIleu	140
Db	361	CGCTGTTTCTATGTGCTTTTCAAAATCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG	420
Qy	141	LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu	160
Db	421	CTGATGACACTTACATCAGAGATCCCAAGAAAGGGAAATTTTATTTAATGCAATTTGAA	480
Qy	161	ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys	180
Db	481	ACCATGCCCTATGTTAAGAAAAAAGCAGATTGGGCCCTTGGCATCGATAGCAGATAGAAA	540
Qy	181	SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly	200
Db	541	TCTACTTTTGGGAAAGAGTGGTGGCTTTTCCTGCTGTAGAAGAGATTTTCTTCTCAGGA	600

QY 201 SerPheAlaAlaIlePheThrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 DB 601 TCTTTTGTCTGCTATATCTGGCTAAAGAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
 QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 DB 661 ATTGAACCTCATCAGCAGAGATGAAGGACTTCATCTGTGACTTTTGCCTGCTGATGTTCCAA 720
 QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
 DB 721 TACTTAGTAATAAGCCTTCAGAAAGAGAGGTTCAGGAGATCATTTGTGATGCTGTCAA 780
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 DB 781 ATTGACGAGAGTCTTTTAAACAGAGCCTTGCCAGTGGCCTCATTTGGAATGAATTCATT 840
 QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 DB 841 TTGATGAACAGTACATTTGAGTTGTAGCTGCAGAGATTACTTGTGGAACCTTGGATTCTCA 900
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 DB 901 AGGTTTTTCAGGCAGAGAAATCCTTTTGTATTTATGGAAACATTTCTTTAGAGGAAA 960
 QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 DB 961 ACAAAATTTCTTTCAGAAACAGTTCAGAGTATCAGCGTTTTCAGTTTATGGCAGAAACC 1020
 QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 DB 1021 ACAGATAACGCTCTTCACCTTGGATGCAGATTTT 1053
 RESULT 3
 BD064764
 LOCUS Novel protein and DNA thereof. 1053 bp DNA linear PAT 27-AUG-2002
 DEFINITION
 ACCESSION BD064764
 VERSION BD064764.1 GI:22610367
 KEYWORDS JP 2001269184-A/1.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Nakamura,Y., Arakawa,H. and Tanaka,H.
 1 (bases 1 to 1053)
 REFERENCE
 AUTHORS
 TITLE Novel protein and DNA thereof
 JOURNAL Patent: JP 2001269184-A 1 02-OCT-2001;
 YUSUKE NAKAMURA,TAKEDA CHEMICAL INDUSTRIES LTD
 COMMENT OS Homo sapiens (human)
 PN JP 2001269184-A/1
 FD 02-OCT-2001
 PF 27-JUN-2000 JP 2000192401
 PI YUSUKE NAKAMURA,HIROFUMI ARAKAWA,HIROSHI TANAKA PC
 C12N15/09,A61K31/711,A61K38/00,A61K38/43,A61K45/00,A61K48/00, PC
 A61P35/00,
 PC A61P43/00,C07K16/32,C07K16/40,C12N1/15,C12N1/19,C12N1/21 PC
 ,C12N5/10,C12N9/02,
 PC C12P21/02,C12P21/08,G01N33/15,G01N33/50,G01N33/53/(C12N15/09,
 PC C12R1:91),
 PC (C12N15/09,C12R1:19),(C12N9/02,C12R1:91),(C12P21/02,C12R1:91),
 PC (C12P21/08,C12R1:91),C12N15/00,A61K37/02,A61K37/48,C12N5/00,
 PC (C12N15/00,C12R1:91),(C12N15/00,C12R1:19)
 CC Novel protein and DNA thereof
 FH Key Location/Qualifiers
 FT source 1.1053
 /organism='Homo sapiens (human)'.
 FT Location/Qualifiers
 1.1053
 /organism='Homo sapiens'
 /mol_type='genomic DNA'
 /db_xref='taxon:9606'
 FEATURES
 source
 ORIGIN

Alignment Scores:
 Pred. No.: 1,01e-170 Length: 1053
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-10-698-228-1 (1-351) x BD064764 (1-1053)
 QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
 DB 1 ATGGCGCACCCGGAAGCGCGAGCGCGGCTGGATCAGATGAGAGATCAITCTCA 60
 QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
 DB 61 GACACCAACGNAAGTGAATAAAGTCAATAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
 QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
 DB 121 CGGTTTGTCTATCTTCCAATCCAGTACCTCATATTTGAAAAATGTATAAACAGGCACAG 180
 QY 61 AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrPheAsnLys 80
 DB 181 GCTTCCCTCTGGACAGAGAGGTGACTTATCAAGAGATCTCCCTCCTCAGTGAACAG 240
 QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp 100
 DB 241 CTTAAGACAGATGAGAAGTACTTCTCATCTCACATCTTAGCTTTTTCAGCCAGTGTAT 300
 QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 DB 301 GGAATCTTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGACAGGCTCCAGAGGCT 360
 QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 DB 361 CGCTGTTTCTATGGCTTTTCAAAATTTCTATCAGAGAATGTTTCACTCAGAGATGTACAGTTTG 420
 QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 DB 421 CTGATACACACTTATACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
 QY 161 ThrMetProTyrValLysLysLysAlaAspTyrAlaLeuArgTyrIleAlaAspArgLys 180
 DB 481 ACCATGCCCTATGTTAAGAAAAAGACAGATTGGCGCTTCGATGATGATAGATAGAAA 540
 QY 181 SerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePheSerGly 200
 DB 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTCTGCTGTAGAGAGGAGTTTCTTCTCAGGA 600
 QY 201 SerPheAlaAlaIlePheThrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 DB 601 TCTTTTCTGCTATATTTCTGGCTAAAGAAAGAGAGGCTTATGCCAGGACTCACTTTTTC 660
 QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 DB 661 AATGAACCTATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTGCTGATGTTCCAA 720
 QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
 DB 721 TACTTAGTAATAAGCCTTCAGAAAGAGGCTCAGGAGAGATCATTTGTGATGCTGTCAA 780
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 DB 781 ATTGACGAGGAGTCTTTTAAACAGAGCCTTGCAGTTGGCCTCATTTGGAATGAATTCATT 840
 QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 DB 841 TTGATGAACAGTACATTTGAGTTGTAGCTGCAGAGATTACTTGTGGAACCTTGGATTCTCA 900
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 DB 901 AAGGTTTTTCAGGCAGAAAAATCCTTTTGTATTTTATGGAAAAACATTTCTTTAGAGGAAA 960

JOURNAL	Patent: WO 0100799-A 1 04-JAN-2001;	
	TAKEDA CHEMICAL INDUSTRIES LTD, YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA	
COMMENT	OS Homo sapiens (human)	
	PN	WO 0100799-A/1
	PD	04-JAN-2001
	PF	27-JUN-2000 WO 2000JP004189
	PR	28-JUN-1999 JP 99P 181131, 06-JUL-1999 JP 99P 192391 PR
	PI	21-JAN-2000 JP 00P 017770
	PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC	
	C12N9/04, C12N15/53, C12N1/15, C12N1/19, C12N1/21, C12N5/10 PC	
	C12P21/02, A61K38/44	
	PC	A61K48/00, C07K16/40, G01N33/53, C12Q1/26
	CC	
FEATURES	FH	Key Location/Qualifiers.
source	1. 1053	
	/organism="Homo sapiens"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:9606"	
ORIGIN		
Alignment Scores:		
Pred. No.:	1,01e-170	Length: 1053
Score:	1821.00	Matches: 351
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	6	Gaps: 0
US-10-698-228-1 (1-351) x BD093077 (1-1053)		
QY	1	MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSerSer 20
DB	1	ATGGGACCCGGAAGCGCGAGAGCGCGCGGTGGATCAGATGAGAGATCATCTTCA 60
QY	21	AspThrAsnGluSerGluLeuLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
DB	61	GACACCAACGAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
QY	41	ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
DB	121	CGGTTTGTCTATCTTCCATCAGTACCTCGATATTGGAATAATGTATAAACAGGCACAG 180
QY	61	AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
DB	181	GCTTCTCTTCGGACGACGAGAGGTGCACTTATCAAGAGTCTCCCTCCTCGAACCAAG 240
QY	81	LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
DB	241	CTTAAAGCAGATGAGAAGTACTTCATCTCTACATCTTAGGCTTTTTCAGCCAGTGAT 300
QY	101	GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB	301	GGAATTGTAAATGAATAATTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCAGAGGCT 360
QY	121	ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB	361	CGCTGTTCTATGGCTTTTCAAAATCTCATCGAAGATGTTCACTCAGAGATGACAGTTTG 420
QY	141	LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAAATGCAATTGAA 480
QY	161	ThrMetProTyrValLysLysAlaAspTyrAlaLeuArgTyrIleAlaAspArgLys 180
DB	481	ACCATGCCCTATGTTAAGAAAAAAGCAGATTGGGCTTCGCGATGATAGCAGATAGAAA 540
QY	181	SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGluValPhePheSerGly 200
DB	541	TCTACTTTTGGGGAAGATGGTGGCTTTGCTGCTGTAGAAGGAGTTTCTTCTCAGGA 600
QY	201	SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220

DB: 6 Gaps: 0
US-10-698-228-1 (1-351) x BD093086 (1-1053)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
DB 1 ATGGGGGACCCGGAAAGCGCGAGCGCGGGCTGGATCAGGATGAGAGATCATCTTCA 60
QY 21 AspThrAsnGluSerGluLeuLysSerAsnGluProLeuLeuArgLysSerSerArg 40
DB 61 GACACCAACGAAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTCTTCGC 120
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrPheMetTyrLysGlnAlaGln 60
DB 121 CGGTTTGTTCATCTTCCAAATCCAGTACCTGATATTTGGAAATGTATAACAGGACAG 180
QY 61 AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
DB 181 GCTTCTCTCTGACAGCAGAGAGGTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 240
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
DB 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTCTTAGCCCTTTTTCGAGCCAGTGAT 300
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 301 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 361 CGCTGTTTCTATGGCTTTTCAAAATTCATCGAGAGTGTTCACCTCAGAGATGACAGTTG 420
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 421 CTGATAGACACTTACATCAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 480
QY 161 ThrMetProTyrValLysLysLysAlaAspThrAlaLeuArgTyrPheAlaAspArgLys 180
DB 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGATGGATAGCAGATAGAAA 540
QY 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
DB 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGTAGAGAGAGTTCCTTCTCAGGA 600
QY 201 SerPheAlaAlaIlePheThrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 601 TCTTTTGTGCTATATCTTGGCTTAAGAGAGAGAGTCTTATGCCAGGACTCATTTCCTCC 660
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
DB 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
DB 721 TACTTAGTAAATTAAGCTTTCAGAAAGAGGGTCCAGGAGATCATTTGTATGCTGTCAA 780
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
DB 781 ATGTAGCAGAGAGTTTTAAACAGAGCTTGGCCAGTTGGCTCATTTGGAAATGAATTCATT 840
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
DB 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATTTGTGGAACCTTGGATTCCTCA 900
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
DB 901 AAGGTTTTTTCAGGCGAAAAATCCTTTGATTTTATGAAAAACATTTCTTTAGAAAGGAAA 960
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
DB 961 ACAAATTTCTTTGAGAAACAGTTCCTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC 1020
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

DB 1021 ACAGATAACGTCTTCCCTTGGATGCAGATTTT 1053

RESULT 7
CQ714252
LOCUS CQ714252 1056 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 186 from Patent WO02068579.
ACCESSION CQ714252
VERSION CQ714252.1 GI:42275109
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 186 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
Location/Qualifiers
source 1..1056
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 1,01e-170 Length: 1056
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-698-228-1 (1-351) x CQ714252 (1-1056)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
DB 1 ATGGGGGACCCGGAAAGCGCGAGCGCGGGCTGGATCAGGATGAGAGATCATCTTCA 60
QY 21 AspThrAsnGluSerGluLysSerAsnGluProLeuLeuArgLysSerSerArg 40
DB 61 GACACCAACGAAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCCTGC 120
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrPheMetTyrLysGlnAlaGln 60
DB 121 CGGTTTGTTCATCTTCCAAATTCAGTACCTGATATTTGGAAATGTATAACAGGACAG 180
QY 61 AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
DB 181 GCTTCTCTTTCGACAGCAGAGAGGTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 240
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
DB 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTCTTAGCCCTTTTTCGAGCCAGTGAT 300
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 301 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 361 CGCTGTTTCTATGGCTTTTCAAAATTCATCGAGAGTGTTCACCTCAGAGATGACAGTTG 420
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 421 CTGATAGACACTTACATCAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 480
QY 161 ThrMetProTyrValLysLysLysAlaAspThrAlaLeuArgTyrPheAlaAspArgLys 180
DB 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGATGGATAGCAGATAGAAA 540

```

Qy 181 SerThrPheGlyGluArgValAlaPheAlaValGluGlyValPhePheSerGly 200
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCTTGTCTGCTAGAGGAGTTTCTTCTCAGGA 600

Qy 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 601 TCTTTTGTCTCTATATTCTCGCTAAAGAGAGGCTTATGCCAGGACTCACATTTTTC 660

Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 661 AATGAACATCATCAGCAGATGAAGGACTTCACTGTGACTTTGCTTGTCTGAGTTTCAA 720

Qy 241 TyrLeuValAsnLysProSerGluGluArgValArgGluLeuIleValAspAlaValLys 260
Db 721 TACTAGTAATAAGCCCTCAGAAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAA 780

Qy 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 781 ATTGAGCAGGAGTTTAAACAGAGCCCTGCGAGTTGCGCTCATTTGGAATGAATTGCATT 840

Qy 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 841 TTGATGAAACAGTACATGAGTTGTTGAGCTGACAGATTTACTTTGGAACTTTGGATTCTCA 900

Qy 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 901 AAGTTTTCAGGCAGCAAAATCCTTTTGAATTTATGGAACAATTTCTTTAGAGGAAA 960

Qy 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 961 ACAATTTCTTTGAGAAACAGATTTTCAAGATATCAGCGTTTTCAGTTTATGGCAGAAACC 1020

Qy 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1021 ACAGATAACGCTCTTACCTGGTATGCAGATTTT 1053

RESULT 8
LOCUS AB166671
DEFINITION Homo sapiens p53R2 mRNA for p53-inducible ribonucleotide reductase
small subunit 2, complete cds.
ACCESSION AB166671
VERSION AB166671.1 GI:45259568
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
MAMMALIA; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 Ugai,H. and Yokoyama,K.K.
Homo sapiens p53R2 mRNA for p53-inducible ribonucleotide reductase
small subunit 2, complete cds
Published Only in Database (2004)
JOURNAL
REFERENCE
2 (bases 1 to 1056)
AUTHORS Ugai,H. and Yokoyama,K.K.
TITLE Direct Submission
JOURNAL
SUBMITTED (05-MAR-2004) Hideyo Ugai, RIKEN BioResource Center, Gene
Engineering Division; 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074,
Japan (E-mail:ugai@rtc.riken.go.jp, Tel:81-29-836-3612,
Fax:81-29-836-9120)
FEATURES
source
1..1056
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="p53R2_C19"
/cell_line="Saos-2"
/cell_type="Osteogenic sarcoma"
1..1056
/gene="p53R2"
1..1056
/gene="p53R2"

```

```

/codon_start=1
/product="p53-inducible ribonucleotide reductase small
subunit 2"
/protein_id="BAD12267.1"
/db_xref="GI:45259569"
/translation="MGDPEREAAAGLDODERSSDTNESEIKSNRPLLRKSSRRFVI
FPDPIDIKMYKQAFWTAEBVDLKDLPWNKLKADKXFISHILAFPAASDGI
VNEKLVERFQEVQVPEARCFYQQLIENHSEMYSLIDITYRDPKKEFLFNAIE
TMPYVKKADWALRWIADKSTFGRVFAFAVEGFFSGFAAI FWLKKRGLMPGIT
PSNLIISDEGLHCDPACILFOYLVPKSEERVREIIVDAVKIEQSLFLEALPVLIG
MNCILMKQYIEFVADRLVLELGSFKVQFAENPFDFMENISLEGKTFFFEKRVSEYQRF
AVMAETDNDVFTLDAP"
ORIGIN
Alignment Scores:
Pred. No.: 1,01e-170 Length: 1056
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-10-698-228-1 (1-351) x AB166671 (1-1056)
Qy 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20
Db 1 ATGGCGGACCCGGAAGGCGGAGCGCGGCTGGATCAGATGAGATCATCTTCA 60
Qy 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
Db 61 GACACCAACGAAAGTGAATAAAGTCAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Qy 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
Db 121 CGTTTGTCTCTTTTCCAATCCAGTACCTGTATTTGGAAATATGATATAACAGGCACAG 180
Qy 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
Db 181 GCTTCTTCTGGACAGCAGAGAGGTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 240
Qy 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Db 241 CTTAAAGCAGATGAGAAGTACTTCTCATCATCTTACCTTTAGCCCTTTTTCAGCCAGTGAT 300
Qy 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 301 GGAATTGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGTTCAGAGGCT 360
Qy 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAATGTTTCACTCAGAGATGTACAGTTTG 420
Qy 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAAATCAATGAA 480
Qy 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Db 481 ACCATGCCCTATGTTAAAGAAAAAGCAGATGCGGCTTCGATGGATAGCAGATAGAAAA 540
Qy 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTAGAGGAGTTTCTTCTCAGGA 600
Qy 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 601 TCTTTTGTCTATATTCTGGCTAAAGAGAGAGGCTTATGCGAGGACTCACTTTTTC 660
Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 661 AATGAACATCATCAGCAGATGAAGGACTTCTCATCTGACTTTGCTTGTGCTGATTTCCAA 720
Qy 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleValAspAlaValLys 260

```


PI YUSUKE NAKAMURA,HIROFUMI ARAKAWA,HIROSHI TANAKA PC
 C12N15/09,A61K31/711,A61K38/00,A61K38/43,A61K45/00,A61K48/00, PC
 A61P35/00,
 PC A61P43/00,C07K16/32,C07K16/40,C12N1/15,C12N1/19,C12N1/21 PC
 C12N5/10,C12N9/02,
 PC C12P21/02,C12P21/08,G01N33/15,G01N33/50,G01N33/53// (C12N15/09,
 C12R1/91),
 PC (C12N15/09,C12R1/91), (C12N9/02,C12R1/91), (C12P21/02,C12R1/91),
 PC (C12P21/08,C12R1/91), (C12N15/00,A61K37/02,A61K37/48,C12N5/00,
 C12N15/00,C12R1/91), (C12N15/00,C12R1/91)
 CC Novel protein and DNA thereof
 FH Key Location/Qualifiers
 FT source 1..1081 /organism="Homo sapiens (human)".
 FEATURES
 source
 1..1081 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

ORIGIN
 Alignment Scores:
 Pred. No.: 1 04e-170 Length: 1081
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-698-228-1 (1-351) x BD064766 (1-1081)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
 DB 20 ATGGGCGACCGAAGCGCGAAGCGCGCGCTGGATCAGATGAGATCATCTTCA 79
 QY 21 AspThrAsnGluSerGluLeuLysSerAsnGluGluProLeuLeuArgLysSerArg 40
 DB 80 GACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTCAAGAAAGAGTTCTCGC 139
 QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
 DB 140 CGGTGTTGTCACTTCCATCCAGTACCTCCCTGATATTGGAAATGATATAACAGGCACAG 199
 QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
 DB 200 GCTTCCTCTTGGACAGCAGAGAGGTGACTTATCAAGAGATCTCCCTCACTGGAACAG 259
 QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaIleAsp 100
 DB 260 CTTTAAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTAGCCCTTTTTCGACGCAGTGAT 319
 QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 DB 320 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCGAGGTTCAGAGGCT 379
 QY 121 ArgCysPheTrpGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 DB 380 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAATGTTCCACTCAGAGATGTACAGTTG 439
 QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 DB 440 CTGATAGACACTTACATCAGATGCCAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 499
 QY 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
 DB 500 ACCATGCCCTATGTTAAGAAAAAAGCAGATTGGGCTTGGATGGATAGCAGATAGAAAA 559
 QY 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluValPhePheSerGly 200
 DB 560 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAAGAGAGTTTCTTCTCAGGA 619
 QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220

DB 620 TCTTTTGTCTATATTCTGGCTAAGAGAGAGGTCTTATGCCAGGACTCACTTTTCC 679
 QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 DB 680 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTCCAA 739
 QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
 DB 740 TACTTAGTAATAAGCCCTTCAGAAAGAGGTTCAGGAGATCATTTGTGATGCTGTCAA 799
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 DB 800 ATTGACGAGGAGTTTTHAACAGAGCCCTTGCAGTTGGCCTCATTTGGAATGAATTCATT 859
 QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuGlyPheSer 300
 DB 860 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGATTTCTCA 919
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGlyLys 320
 DB 920 AAGGTTTTTCAGGCAGAAAAATCCTTTTGATTTTATGGAAAAATTTCTTTTAGAGGAAAA 979
 QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 DB 980 ACAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTTCAGTTATGCGAAACC 1039
 QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 DB 1040 ACAGATAACGTTCTTCACTTGGATGCAGATTTT 1072

RESULT 11
 BD093079
 LOCUS Novel protein and its DNA.
 DEFINITION
 ACCESSION BD093079
 VERSION BD093079.1 Gi:22638667
 KEYWORDS WO 0100799-A/3.
 SOURCE Homo sapiens (human)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Nakamura,Y., Arakawa,H. and Tanaka,H.
 Novel protein and its DNA
 Patent: WO 0100799-A 3 04-JAN-2001;
 TAKEDA CHEMICAL INDUSTRIES LTD,YOSUKE NAKAMURA,HIROFUMI ARAKAWA,
 HIROSHI TANAKA
 COMMENT
 OS Homo sapiens (human)
 PN WO 0100799-A/3
 PD 04-JAN-2001
 PF 27-JUN-2000 WO 2000JP004189
 PR 28-JUN-1999 JP 99P 181131,06-JUL-1999 JP 99P 192391 PR
 21-JAN-2000 JP 00P 017770
 PI YUSUKE NAKAMURA,HIROFUMI ARAKAWA,HIROSHI TANAKA PC
 C12N9/04,C12N15/53,C12N1/15,C12N1/19,C12N1/21,C12N5/10 PC
 C12P21/02,A61K38/44,
 PC A61K48/00,C07K16/40,G01N33/53,C12Q1/26
 CC

FEATURES
 FH Key Location/Qualifiers.
 1..1081
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

ORIGIN
 Alignment Scores:
 Pred. No.: 1 04e-170 Length: 1081
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 Db 845 TCTTTTGTCTCTATATTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACATTTTTC 904
 QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 Db 905 AATGAACTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA 964
 QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluLeuIleValAspAlaValLys 260
 Db 965 TACTTAGTAATAAGCCCTCAGAGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAA 1024
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 Db 1025 ATTGAGCAGGAGTATTTAAACAGAGCCCTTGCAGTTCGCTCATTTGGAATGAATTGCATT 1084
 QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 Db 1085 TTGATGAACAGTACATGATGTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTTCTCA 1144
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 Db 1145 AAGTATTTTTCAGGAGAAATCCTTTTGATTTTATGGAACATTTCTTTAGAGGAAA 1204
 QY 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 Db 1205 ACAAATTTCTTTGAGAAACAGATTTTCAGATATCAGCGTTTTCAGTTATGGCAGAAACC 1264
 QY 341 ThrAsnValPheThrLeuAspAlaAspPhe 351
 Db 1265 ACAGATAACGCTTTCACCTGGTATGCAGATTTT 1297

RESULT 13

BD064765
 LOCUS BD064765
 DEFINITION Novel protein and DNA thereof.
 ACCESSION BD064765
 VERSION BD064765.1 GI:22610368
 KEYWORDS JP 2001269184-A/2.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 4955)
 Nakamura, Y., Arakawa, H. and Tanaka, H.
 Novel protein and DNA thereof
 Patent: JP 2001269184-A 2 02-OCT-2001;
 YUSUKE NAKAMURA, TAKEDA CHEMICAL INDUSTRIES LTD
 OS Homo sapiens (human)
 PN JP 2001269184-A/2
 PD 02-OCT-2001
 PF 27-JUN-2000 JP 2000192401
 PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
 C12N15/09, A61K31/711, A61K38/00, A61K45/00, A61K48/00, PC
 A61P35/00,
 PC A61P43/00, C07K16/32, C07K16/40, C12N1/15, C12N1/19, C12N1/21 PC
 , C12N5/10, C12N9/02,
 PC C12P21/02, C12P21/08, G01N33/15, G01N33/50, G01N33/53// (C12N15/09,
 PC C12R1:91),
 PC (C12N15/09, C12R1:19), (C12N9/02, C12R1:91), (C12P21/02, C12R1:91),
 PC (C12P21/08, C12R1:91), C12N15/00, A61K37/02, A61K37/48, C12N5/00,
 PC (C12N15/00, C12R1:91), (C12N15/00, C12R1:19)
 CC Novel protein and DNA thereof
 FH Key Location/Qualifiers
 FT source 1..4955
 FT /organism='Homo sapiens (human)'.
 FT Location/Qualifiers
 1..4955
 /organism='Homo sapiens'
 /mol_type='genomic DNA'
 /db_xref='taxon:9606'

FEATURES
source

ORIGIN

Alignment Scores:
 Pred. No.: 7,39e-170 Length: 4955
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-10-698-228-1 (1-351) x BD064765 (1-4955)
 QY 1 MetGlyAspProGluArgTrpGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20
 Db 245 ATGGCGACCCCGAAAGGCCGAGCGCGGCTGGATCAGATGAGATGATCACTTCA 304
 QY 21 AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerSerArg 40
 Db 305 GACACCAACGAAGTGAATAAAGTCAATGAAGGCCACTCTTAAGAAAGAGTTCTCGC 364
 QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
 Db 365 CGGTTTGTCTCTTCCAAATCCAGTACCTGTATATTGGAAAATGTATTAACAGGCACAG 424
 QY 61 AlaSerPheTrpThrAlaGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
 Db 425 GCCTTCCTTCTGGACAGCAGAGAGTGCACCTTATCAAGGATCTCCCTCACTGGAACAAG 484
 QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
 Db 485 CTTAAGCAGATGAGAGTACTTCACTCTCTCATCTTAGCCCTTTTTCAGCCAGTGAT 544
 QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 Db 545 GGAATTTGTAATGAAATTTGTGTGAGCGCTTTAGTCAGGAGGTGCAGGTTCAGAGGCT 604
 QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 Db 605 CGCTGTTTCTATGGCTTTTCAAAATTTCTATCGAAGATGTTCTACTCAGAGATGTACAGTTG 664
 QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA 724
 QY 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
 Db 725 ACCATGCCCTATGTAAGAAAAACAGATGGGCTTGGGATGATGATGATGATGATGATA 784
 QY 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
 Db 785 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGTCTGTAGAGGAGTCTTCTTCTCAGGA 844
 QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 Db 845 TCTTTTGTCTGTATATTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 904
 QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 Db 905 AATGAACTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA 964
 QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
 Db 965 TACTTAGTAATAAGCCCTTCAAGAAAGGCTCAGGAGATCATTTGTGATGTCGTCAAA 1024
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 Db 1025 ATTGAGCAGGAGTATTTAAACAGAGCCCTTGCAGTTCGCTCATTTGGAATGAATTGCATT 1084
 QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 Db 1085 TTGATGAACAGTACATGATGTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTTCTCA 1144
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 Db 1145 AAGTATTTTTCAGGAGAAATCCTTTTGATTTTATGGAACATTTCTTTTAGAAGGAAA 1204

```

QY 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 1205 ACAAAATTTCTTTGAGAAACAGGTTTCAGAGATATCAGCGTTTGTGAGTTATGGCAGAAACC 1264

QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1265 ACAGATAACGTCCTACCTCGGATGAGATTTT 1297

RESULT 14
BD093078 4955 bp DNA linear PAT 27-AUG-2002
LOCUS Novel protein and its DNA.
DEFINITION BD093078
ACCESSION BD093078
VERSION WO 0100799-A/2.
KEYWORDS Novel protein and its DNA
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4955)
AUTHORS Nakamura, Y., Arakawa, H. and Tanaka, H.
TITLE Novel protein and its DNA
JOURNAL Patent: WO 0100799-A 2 04-JAN-2001.
TAKEDA CHEMICAL INDUSTRIES LTD, YUSUKE NAKAMURA, HIROFUMI ARAKAWA,
HIROSHI TANAKA
COMMENT OS Homo sapiens (human)
PN WO 0100799-A/2
PD 04-JAN-2001
PF 27-JUN-2000 WO 2000JP004189
PR 28-JUN-1999 JP 99P 181131, 06-JUL-1999 JP 99P 192391 PR
21-JAN-2000 JP 00P 017770
PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
C12N9/04, C12N15/53, C12N1/15, C12N1/19, C12N1/21, C12N5/10 PC
.C12P21/02, A61K38/44.
PC A61K48/00, C07K16/40, G01N33/53, C12Q1/26
CC
FH Key Location/Qualifiers.
FEATURES
source
1. .4955
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred: No.: 7 39e-170 Length: 4955
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-698-228-1 (1-351) x BD093078 (1-4955)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
Db 245 ATGGGGGACCCGAAAGCGCGAAGCGCGGCTGGATCAGATCAGAGATCATCTTCA 304

QY 21 AspThrAsnGluSerGluLeuLysSerAsnGluGluProLeuLeuArgLysSerArg 40
Db 305 GACACCAACGAAAGTAAATGAAATGAAATGAAAGAGCCATCTCTAAGAAAGAGTTCTCGC 364

QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
Db 365 CGGTGTGTCTCTTCCATTCAGTACCTGACCTGATATTTGGAAATGATATAACAGGCACAG 424

QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
Db 425 GTTCTCTCTGGACACGACAGAGGTCGACTATCAAGGAATCTCCTCCTCTGGAACACAG 484

QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp 100

```

```

Db 485 CTTAAACGACATGAGAAAGTACTTCTCATCTCTCACATCTTAGCCCTTTTTCAGCCAGTGAT 544
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 545 GGAATTTGAAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 604

QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 605 CGCTGTTTCTATGGCTTTTCAAAATTTCTCATCGAATGTTTCACTCAGAGATGTACAGTTG 664

QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 724

QY 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Db 725 ACCATGCCCTATGTTAAGAAAAAGCAGATGGGCCCTTGCATGATGATAGCAGATAGAAA 784

QY 181 SerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePheSerGly 200
Db 785 TCTACTTTTGGGAAAGAGTGGTGGCTTGTCTGCTGTAGAAAGAGTTTCTTCTCAGGA 844

QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 845 TCITTGTGCTCTATATTTCTGGCTAAGAGAGAGGTCTTATGCCAGGACTCATTTTTCC 904

QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 905 AATGAACCTCATCAGCAGATGAAGGACTTCACTGTGCTGCTGCTGCTGCTGCTGCTGCT 964

QY 241 TyrIleuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
Db 965 TACTTAGTAAATAGCCCTTCAGAAAGAGGTTCAGGAGAGATCATTTGTGATGCTGTCAAA 1024

QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 1025 ATTGACGAGGAGTTTAAACAGAGCCCTGCCAGTGGCTTCATTTGGAATGAATGCAIT 1084

QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 1085 TTGATGAACAGTACATGAGTTTGTAGTGCAGAGATTACTTGTGGAACCTTGGATTCTCA 1144

QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 1145 AAGGTTTTTCAGGCAGAAAAATCCTTTTGAATTTATGGAATAACATTTCTTTAGAAGGAAA 1204

QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 1205 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGAGTTATGGCAGAAACC 1264

QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1265 ACAGATAACGTCCTTCACTTGGATGAGATTTT 1297

RESULT 15
AB036063 4955 bp mRNA linear PRI 11-MAR-2000
LOCUS Homo sapiens p53R2 mRNA for ribonucleotide reductase, complete cds.
DEFINITION AB036063
ACCESSION AB036063
VERSION AB036063.1 GI:7229085
KEYWORDS p53R2; ribonucleotide reductase.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Tanaka, H., Arakawa, H., Yamaguchi, T., Shiraishi, K., Fukuda, S.,
Matsumi, K., Takei, Y. and Nakamura, Y.
TITLE A ribonucleotide reductase gene involved in a p53-dependent
cell-cycle checkpoint for DNA damage
JOURNAL Nature 404 (6773), 42-49 (2000)
MEDLINE 20179179
PUBMED 10716435

```

REFERENCE 2 (bases 1 to 4955)
 AUTHORS Tanaka,H., Arakawa,H. and Nakamura,Y.
 TITLE Direct Submission
 JOURNAL Submitted (18-DEC-1999) Yusuke Nakamura, University of Tokyo,
 Institute of Medical Science, Human Genome Center, Laboratory of
 Molecular Medicine; 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639,
 Japan (E-mail:yusuke@ims.u-tokyo.ac.jp, Tel:+81-3-5449-5372,
 Fax:+81-3-5449-5433)

FEATURES
 source
 1. .4955
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="8"
 /map="8q23.1"
 gene
 1. .4955
 /gene="p53R2"
 245..1300
 /genes="p53R2"
 /codon_start=1
 /product="ribonucleotide reductase"
 /protein_id="BAA92434.1"
 /db_xref="GI:7229086"
 /translation="MGDPPEPAAGLDQDERSSDITNESBIKSNRPFLKSRFRVFI
 FPIQPDIIWMYKQASFTWAEVDLSKDLPHNKLKADEKFIPISHLAFPAASDGI
 VNENLVERFSQEVQPEARCFYQILNIENVHSEMYSLIDTIRDPKKRGLFNATIE
 TWPYVKKADWALRWLRADRSKSTGRVAFVAVEGFFSGSPAAIIFWLKKRGLMPGLT
 FSNELISRDGHLGDFACILMFOVLVKNKPSSEVRRELIIVDAVKIEQSLFTEALPVGLIG
 MNCILMKQIIEFVADRLVLELGFPSKVFQAEENPDFMENISLEKTNFFPKRVSEYQRF
 AVMAETIDNVITLUDAF"
 4933
 polyA_site
 /gene="p53R2"
 /note="22 a nucleotides"

ORIGIN
 Alignment Scores:
 Pred. No.: 7,39e-170 Length: 4955
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-10-698-228-1 (1-351) x AB036063 (1-4955)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
 DB 245 ATGGGCGACCCGGAAGCGCGGAAGCGCGCGCTGGATCAGATGAGATCATCTTCA 304
 QY 21 AspThrAsnGluSerGluLeuLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
 DB 305 GACACCAACGAAAGTGAATTAAGTCAATGAGAGCCACTCTTAGAAGAGATTCTCGC 364
 QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
 DB 365 CGGTTTGTTCATCTTTCCATCCAGTACCTCGATTTTGGAAATGATATAAACAGGCACAG 424
 QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
 DB 425 GCTTCTCTTGGACAGACAGAGAGGTGCGACTTATCAAGAGATCTCCCTCACTGGAAACAAG 484
 QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
 DB 485 CTTAAAGCAGATGAGAAGTACTTCATCTCTCACTCTTACCTTTTGGCAGCGAGTAT 544
 QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 DB 545 GGAATTGTAATCAAAATTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGGCTTCCAGAGGCT 604
 QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 DB 605 CGCTGTTTCTATGGCTTCAAAATCTCATCAGAAATGTTTCACTCAGAGATGTACAGTTTG 664

QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 DB 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA 724
 QY 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTyrIleAlaAspArgLys 180
 DB 725 ACCATGCCCTATGTTAAGAAAAGACAGATTGGGCTTGGCATGGATGACAGATAGAAA 784
 QY 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
 DB 785 TCTACTTTTGGGAAAGAGTGGTGGCTTGTCTGTAGTAGAGGAGTTTCTTCTCAGGA 844
 QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 DB 845 TCTTTTGTCTGCTATATTCTGGCTAAGAGAGAGAGGTCTTATGCCAGGACTCACTTTTCC 904
 QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 DB 905 ATGAACTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGTGCTGCTGATGTTCCAA 964
 QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
 DB 965 TACTTAGTAATAAGCCTTCAGAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 1024
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 DB 1025 ATTGAGCAGGAGTATTTAAACAGAAAGCCTTGCAGTTGGCCTCATTTGGAATGAATTGCATT 1084
 QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 DB 1085 TTGATGAACAAGTACATTTGAGTTTGTAGCTGACAGATTACTTTGTGGAACCTTGGATTCTCA 1144
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 DB 1145 AAGTTTTTTCAGGCAGAAAATCCTTTTGATTTTATGGAACATTTCTTTAGAGGAAA 1204
 QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 DB 1205 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTTGCAGTTTATGCGAGAAACC 1264
 QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 DB 1265 ACAGATAACGCTCTTCCCTTGGATGCGAGATTTT 1297

Search completed: June 16, 2005, 13:06:51

Job time : 5183 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 16, 2005, 17:22:59 ; Search time 213 Seconds
(without alignments)
2696.400 Million cell updates/sec

Title: US-10-698-228-1

Perfect score: 1821
Sequence: 1 MGDPERPEAGLDQDERSS.....QRFVNAETDNNVFTLDADF 351

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DRV=xlh
-Q=/cgn2_1/USPTO.spool/US10698228/runat_14062005_161417_28327/app.query.fasta_1.519
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10698228@cgn_1_177@runat_14062005_161417_28327 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1821	100.0	1053	US-10-019-733-2	Sequence 2, Appli
2	1821	100.0	1053	US-10-019-733-12	Sequence 12, Appli
3	1821	100.0	1081	US-10-019-733-4	Sequence 4, Appli
4	1821	100.0	4955	US-10-019-733-3	Sequence 3, Appli
5	1460.5	80.2	2479	US-09-949-016-2025	Sequence 2025, Ap
6	1460.5	80.2	2500	US-09-962-665-9	Sequence 9, Appli
7	1460.5	80.2	2500	US-09-923-655-1370	Sequence 1370, Ap
8	1460.5	80.2	2500	US-09-963-333-9	Sequence 9, Appli
9	1460.5	80.2	2500	US-09-962-677-9	Sequence 9, Appli
10	1460.5	80.2	2500	US-09-949-016-145	Sequence 145, App
11	1210	66.4	14176	US-08-307-499-1	Sequence 1, Appli
12	1210	66.4	14176	US-08-307-499-14	Sequence 14, Appli

c

c	13	1210	66.4	14176	3	US-09-299-268-1	Sequence 1, Appli
	14	1210	66.4	14176	3	US-09-299-268-14	Sequence 14, Appl
	15	1131	62.1	1236	4	US-09-248-796A-3914	Sequence 3914, Ap
	16	1072.5	58.9	825	1	US-08-307-499-40	Sequence 40, Appl
	17	1072.5	58.9	825	3	US-09-299-268-40	Sequence 40, Appl
	18	984.5	54.1	1112	1	US-08-136-743B-1	Sequence 1, Appli
	19	969.5	53.2	1083	4	US-09-248-796A-3915	Sequence 3915, Ap
	20	768	42.2	481	3	US-08-905-223-125	Sequence 125, App
	21	674	37.0	656	3	US-09-328-111-791	Sequence 791, App
	22	599	32.9	449	4	US-09-270-767-14670	Sequence 14670, A
	23	558	30.6	419	3	US-08-905-223-129	Sequence 129, App
	24	543.5	29.8	11820	4	US-09-949-016-11887	Sequence 11887, A
	25	543.5	29.8	11826	4	US-09-949-016-13767	Sequence 13767, A
c	26	513.5	28.0	13706	4	US-09-902-540-1124	Sequence 1124, Ap
	27	509.5	28.0	1053	4	US-09-902-540-2666	Sequence 2666, Ap
	28	488.5	26.8	366	1	US-08-307-499-13	Sequence 13, Appl
	29	488.5	26.8	366	3	US-09-299-268-13	Sequence 4435, Ap
	30	322.5	21.6	297	4	US-09-313-294A-4435	Sequence 588, App
	31	381.5	21.0	1338	4	US-09-328-352-588	Sequence 3180, App
c	32	378.5	20.8	1707	4	US-09-252-991A-3180	Sequence 2787, Ap
	33	377.5	20.7	1431	4	US-09-252-991A-2787	Sequence 2692, Ap
	34	370	20.3	984	4	US-09-252-991A-2692	Sequence 319, App
	35	355	19.5	373	4	US-09-401-064-319	Sequence 337, App
c	36	355	19.5	373	4	US-09-401-064-337	Sequence 1, Appli
	37	338	18.6	3278	1	US-07-778-890A-1	Sequence 1, Appli
	38	335	18.4	1230025	4	US-09-198-452A-1	Sequence 1, Appli
	39	335	18.4	1230230	4	US-09-438-185A-1	Sequence 76, Appl
	40	330	18.1	124884	4	US-09-661-596A-76	Sequence 1, Appli
	41	330	18.1	124884	4	US-09-913-514-1	Sequence 2, Appli
c	42	330	18.1	125157	4	US-09-913-514-2	Sequence 13, Appl
	43	317	17.4	3678	4	US-09-112-580-13	Sequence 19, Appl
c	44	307	16.9	35100	2	US-08-770-379-19	Sequence 19, Appl
	45	307	16.9	35100	3	US-08-757-669A-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1
US-10-019-733-2
; Sequence 2, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-2

Alignment Scores:	6.06e-227	Length:	1053
Pred. No.:	1821.00	Matches:	351
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	100.00%	Gaps:	0
DB:	4		

US-10-698-228-1 (1-351) x US-10-019-733-2 (1-1053)

Oy 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20

Db	1	ATGGGCGACCCGGAAAGCCCGGAAGCGCGCGGTGGATCAGGATGAGAGATCATCTTCA	60
Qy	21	AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerSerArg	40
Db	61	GACACCACGAAAGTGAATAAAGTCAATGAAGGCCACTCTCTAAGAAAGATTCTCGC	120
Qy	41	ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln	60
Db	121	CGGTTTGTCACTCTTCCAAATCCAGTCCCTGATATTTGGAAAATGTATAAACAGGCACAG	180
Qy	61	AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys	80
Db	181	GCCTCTCTTGGACAGCAGAAAGAGGTGGACTATTCAAAGGATCTCCCTCATCTGGAACAAG	240
Qy	81	LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp	100
Db	241	CTTAAACAGATGAGAAGTACTTCATCTCTCACATCTTAGCGCTTTTGGAGCCAGTGAT	300
Qy	101	GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla	120
Db	301	GGAAATTTGTAATGAAAAATTTGGTGGAGCGCTTTAAGTCAGGAGGTGCGAGTTCCAGAGCT	360
Qy	121	ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu	140
Db	361	CGCTGTCTTATGGCTTTCAAATTTCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG	420
Qy	141	LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu	160
Db	421	CTGATACACACTTACATCAGAGATCCCAAGAAAAGGGAAATTTTATTATTAAATGCAATTGAA	480
Qy	161	ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys	180
Db	481	ACCATGCCCTATGTTTAAAGAAAAAGCAGATTGGGCTTGGCATGGATAGCAGATAGAAAA	540
Qy	181	SerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePheSerGly	200
Db	541	TCCTACTTTGGGGAAGAGTGGTGGCCCTTGCTGCTGTAGAAGGAGATTCTCTCTCAGGA	600
Qy	201	SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer	220
Db	601	TCITTTTGCTGCTATATTTCTGGCTAAAGACAGAGAGGTCTTATGCCAGGACTCACITTTTCC	660
Qy	221	AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln	240
Db	661	AATGAACCTCATCAGCAGAGATGAAGGACTTCACCTGTGACTTTGCTGTGCTGATGTTCCAA	720
Qy	241	TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys	260
Db	721	TACTTTAGTAATAATAGCCCTCAGAAAAGGGTCAAGGAGATCATTTGTGTGCTGTCAAA	780
Qy	261	IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle	280
Db	781	ATTGAGCAGAGATTTTTTAACAGAGCCCTGGCAGTTGGCCTCATTTGGAATGAATTCATT	840
Qy	281	LeuMetLysGlnTyrIleGluPheValAlaAspArgIleuValGluLeuGlyPheSer	300
Db	841	TTGATGAAACAGTACATTTGATGTTGTAGCTGACAGATTACTTTGTGGAACTTTGGATTTCTCA	900
Qy	301	LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys	320
Db	901	AAGTTTTTTCAGGCAGAAAATCTCTTTTGATTTTATGGAACAATTTCTTTATGAAGGAAAA	960
Qy	321	ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr	340
Db	961	ACAAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTTATTCGCGAAACC	1020
Qy	341	ThrAspAsnValPheThrLeuAspAlaAspPhe	351
Db	1021	ACAGATAACGCTCTTTCACCTTGGATGACAGATTTT	1053

RESULT 2

US-10-019-733-12

```

; Sequence 12, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WC00P
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 12
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-12

Alignment Scores:
Pred. No.: 6,06e-227 Length: 1053
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

```

Alignment Scores:

Pred. No.:	6.06-227	Length:	1053
Score:	1821.00	Matches:	351
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-10-698-228-1 (1-351) x US-10-019-733-12 (1-1053)

Qy	1	MetGlyAspProGluAurArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer	20
Db	1	ATGGCGCAACCGGAAAGCCGGAAAGCGCGGGTGGATCAGGATGAGAGATCATCTTTCA	60
Qy	21	AspThrAsnGluSerGluIleLeuSerAsnGluGluProLeuLeuArgLysSerSerArg	40
Db	61	GACACCAACGAAAGTGAATAAAGTCAAAATGAGAGGCCATCTCTAAGAAAGAGTTCTCGC	120
Qy	41	ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln	60
Db	121	CGGTTTGTTCATCTTTCCAAATCCAGTACCCCTCATATTTCGAAAAATGATATAAACACAGGCACAG	180
Qy	61	AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys	80
Db	181	GCITCTCTTCGGACAGCAGAGAGGTTGACTTATCAAAAGGATCTCCCTCCTCTGGAAACAG	240
Qy	81	LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp	100
Db	241	CTTAAAGCAGATGAGAAGTACTTCATCTCTCACATCTTAGCCCTTTTTCGACCCAGTGAAT	300
Qy	101	GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla	120
Db	301	GGAAATCTAAATGAAATTTGGTCGAGCGCTTAGTCAGGAGGTCAGGTTCCAGAGGCT	360
Qy	121	ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu	140
Db	361	CGCTGTGTTCTATGGCTTTCAAATTTCTATCGAGATGTTCTCATCGAGATGTGTACAGTTTG	420
Qy	141	LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu	160
Db	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGAGGGAATTTTATTTAATGCAATTGAA	480
Qy	161	ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys	180
Db	481	ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCCCTTCGATGATAGCAGATAGAAA	540
Qy	181	SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly	200
Db	541	TCTACTTTTGGGAAAGAGTGTGGCCCTTCTGCTGTAGAAGGAGTGTCTTCTTCAGGA	600

QY 201 SerPheAlaAlaIlePheThrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 Db 601 TCTTTTGGCTGCTATATCTGGCTAAAGAGAGAGAGGCTTATGCGAGAGACTCATTCTTTTCC 660
 QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 Db 661 AATGAACCTATCAGCAGAGATGAAGACTTCTCATCTGTGACCTTGTGCTGCTGATGTTCAA 720
 QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleLeuValAspAlaValLys 260
 Db 721 TACTTAGTAATAAGCTTTCAGAGAAAGAGGCTCAGGAGATCATTTGTGATGCTGTCA 780
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 Db 781 ATTGAGCAGGAGTCTTTTAAACAGAGGCTTGCAGTGTGCTCATTTGGAATGAATTCATT 840
 QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 Db 841 TTGATGAACAGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 Db 901 AAGGTTTTTCAGCAGAGAAATCTTTGATTTTATGAAACATTTCTTTAGAGGAAA 960
 QY 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 Db 961 ACAAAATTTCTTTCAGAGAACGAGTTTCAGAGTATCAGGCTTTTGCAGTTATGCGAGAAACC 1020
 QY 341 ThrAsnValPheThrLeuAspAlaAspPhe 351
 Db 1021 ACAGATAACGCTTTCACCTTGGATGCGAGATTTT 1053

RESULT 3

US-10-019-733-4
 ; Sequence 4, Application US/10019733
 ; Patent No. 6682917
 ; GENERAL INFORMATION:
 ; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
 ; TITLE OF INVENTION: New Protein and its DNA
 ; FILE REFERENCE: 2619WOP
 ; CURRENT APPLICATION NUMBER: US/10/019,733
 ; CURRENT FILING DATE: 2001-12-28
 ; PRIOR APPLICATION NUMBER: JP 11-181131
 ; PRIOR FILING DATE: 1999-06-28
 ; PRIOR APPLICATION NUMBER: JP 11-192391
 ; PRIOR FILING DATE: 1999-07-06
 ; PRIOR APPLICATION NUMBER: JP 2000-017770
 ; PRIOR FILING DATE: 2000-01-21
 ; NUMBER OF SEQ ID NOS: 14
 ; SEQ ID NO 4
 ; LENGTH: 1081
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION:
 US-10-019-733-4

Alignment Scores:
 Pred. No.: 6,32e-227 Length: 1081
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-10-698-228-1 (1-351) x US-10-019-733-4 (1-1081)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
 Db 20 ATGGGGACCCGAAAGCGCGAGGCGCGGCTGGATGAGATCATCATCTTCA 79
 QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40

Db 80 GACACCAACGAAAGTCAATTAAGTCAATGAAGAGCCACTCTCTAAGAAAAGAGTTCTCGC 139
 QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrPheMetTyrLysGlnAlaGln 60
 Db 140 CGGTTTGTCTCTTTTCAATCCAGTACCTCATATTTGGAAAATGTATAAACAGGCACAG 199
 QY 61 AlaserPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
 Db 200 GCTTCCCTTCTGACAGCAGAGAGGTTGACTTATCAAGAGGATCTCCCTCACACGGAACAAG 259
 QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
 Db 260 CTTAAAGCAGATGAGAGTACTTCTCATCTCATCTTAGCCTTTTTCACGCCAGTGAT 319
 QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 Db 320 GGAATTGTAAATGAAAATTTGCTGAGGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 379
 QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 Db 380 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCAGAGATGTTCTCATCAGAGATGTACAGTTG 439
 QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 Db 440 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTGAA 499
 QY 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
 Db 500 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGGCTTGGCATGATGATGATGATGATGAT 559
 QY 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGluValPhePheSerGly 200
 Db 560 TCTACTTTGGGAAAGAGTGTGCTTCTGCTGTAGAGAGGAGTTTCTTCTCTCAGGA 619
 QY 201 SerPheAlaAlaIlePheThrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 Db 620 TCTTTTGTGCTATATCTTGGCTAAAGAGAGAGGCTTATGCGAGGACTCATTCTTTTCC 679
 QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 Db 680 AATGAACCTATCAGCAGAGATGAAGACTTCTCATCTGTGACTTGTGCTGCTGATGTTCAA 739
 QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
 Db 740 TACTTAGTAATAAGCTTTCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAAA 799
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 Db 800 ATTGAGCAGGAGTCTTTTAAACAGAGGCTTGCAGTGTGCGCTCATTTGGAATGAATTCATT 859
 QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 Db 860 TTGATGAACAGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 919
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 Db 920 AAGGTTTTTCAGGCGAGAAAATCTTTGATTTTATGAAAACATTTCTTTAGAGGAAA 979
 QY 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 Db 980 ACAAAATTTCTTTCAGAGAACGAGTTTCAGAGTATCAGCGCTTTTTCAGAGTTATGCGAGAAACC 1039
 QY 341 ThrAsnValPheThrLeuAspAlaAspPhe 351
 Db 1040 ACAGATAACGCTTTCACCTTGGATGCGAGATTTT 1072

RESULT 4

US-10-019-733-3
 ; Sequence 3, Application US/10019733
 ; Patent No. 6682917
 ; GENERAL INFORMATION:
 ; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
 ; TITLE OF INVENTION: New Protein and its DNA


```
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 3
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-3

Alignment Scores:
Pred. No.: 7,15e-226 Length: 4955
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-698-228-1 (1-351) x US-10-019-733-3 (1-4955)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
DB 245 ATGGGGCGACCCGGGAAGGCGGAGGCGCGCGGCTGGATGAGATGATGATCTTCA 304
QY 21 AspThrAsnGluSerGluLeuLysSerAsnGluGluProLeuLeuArgLysSerArg 40
DB 305 GACACCAACGAAAGTGAATAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTCTCGC 364
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
DB 365 CGGTTTGTCTATCTTCCATCCAGTACCCTGATATTTGGAAAAATGATTAACAGGCACAG 424
QY 61 AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
DB 425 GCTTCTCTTGGACAGCAAGAGGCGACTTATCAAGAGATCTCCTCTACCTGGAACAG 484
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaLaserAsp 100
DB 485 CTTAAGACGAGTGAAGTACTTCACTCTCACATCTTAGCCCTTTTTCAGCCAGCTGAT 544
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 545 GGAATTGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 604
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 605 CGCTGTTTCTATGGCTTTCAATTTCTCATGAGAATGTTCACTCAGAGATGTACAGTTG 664
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 665 CTGATAGACACTTACATCAGATCCCAAGAAAGGAAATTTTATTTATGCAATTTAA 724
QY 161 ThrMetProTyrValLysLysLysAlaAspTyrAlaLeuArgTyrIleAlaAspArgLys 180
DB 725 ACCATGCCCTATGTTAAGAAAAAGAGATTTGGGCTTGGATGATAGCAGATAGAAAA 784
QY 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
DB 785 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGGAGTCTTCTCTCAGGA 844
QY 201 SerPheAlaAlaIlePheTyrLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 845 TCTTTTGTCTATATCTGGCTAAAGAGAGAGGTCTTATGCCAGACTCACTTTTTC 904
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240

; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/09/949,016
; Sequence 2025, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 2025
; LENGTH: 2479
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-2025

Alignment Scores:
Pred. No.: 2,06e-179 Length: 2479
Score: 1460.50 Matches: 277
Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
DB: 4 Gaps: 1

US-10-698-228-1 (1-351) x US-09-949-016-2025 (1-2479)

QY 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerAspThrAs 23
DB 302 CTTGAGCGGACCCGGTCTCTGGCAGCAAGACCGGAGGAGGATCTTCCAGGAGCCAC 361
QY 23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
DB 362 GGAGCCGAAACTAAAGCAGCTGCCCGCGGTGGAGGATGAGCGCTGTGAGAGAAA 421
QY 38 rSerArgArgPheValIlePhePheProIleGlnTyrProAspIleTyrLysMetTyrLysGl 58
DB 422 CCCCCCGCGTGTGTATCTTCCCATCGAGTACCATGATATCTGCGCAGATGTATAAGAA 481
```

```
QY 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 482 GGCAGAGCGCTCTCTTTGGACCGCGAGAGAGTGGACCTCTCCAGGACATTCAGCACTG 541

QY 78 pAsnLysLeuLysAlaAspGluLysTrpPheLeuSerHisLeuAlaPheAlaAl 98
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 542 GGAATCCCTGAAACCCGAGAGAGATATTTATATCCATGTTCTGGCTTCTTTGAGC 601

QY 98 aSerAspGlyLeuValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 602 AAGCATGGATAGTAATGAAACTTGGTGGAGGATTTAGCCAGAAAGTTTCAGATTAC 661

QY 118 oGluAlaArgCysPheTrpGlyPheGlnIleLeuLeuValHisSerGluMetTy 138
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 662 AGAAGCCCGCTGTTCTATGGCTTCCAAATGGCATGGAACATACATATTCGAAATGTA 721

QY 138 rSerLeuLeuLeuAspThrTrpIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 722 TAGTCTCTTTATTCACACTTACATAAAGATCCCAAGAAAGGAATTTCTCTTCAATGC 781

QY 158 aIleGluThrMetProTrpValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 782 CATGAAACGATGCTTGTGTCAAGAAAGAGGAGACTGGCGCTTGGCTGGATGGGGA 841

QY 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePh 198
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 842 CAAAGAGCTACCTATGTTGAGCTGTTGTAGCTTGTGCTGAGTGAAGCATTTTCTT 901

QY 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 902 TTCGGTCTCTTTGGCTGCGATATTCGCTCAAGAAACGAGGACTGATGCTGCGCTCAC 961

QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 962 ATTTTCTTAATGAACTTATTAGCAGAGATGAGGTTTACACTGTGATTTGCTTGCCTGAT 1021

QY 238 tPheGlnTrpLeuValAsnLysProSerGluGluArgValArgGluIleValAspAl 258
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1022 GTTCAACACCTGGTACACAAACCTCGAGGAGAGTAAAGAAATATTAATCAATGC 1081

QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1082 TGTTGGATAGAACAGAGTTCTCTCACTGAGGCTTGGCTGTGAAGCTCATTTGGATGAA 1141

QY 278 nCysIleLeuMetLysGlnTrpIleGluPheValAlaAspArgLeuLeuValGluLeuGl 298
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1142 TTGCACCTTAATGAAGCAATACATTTGAGTTGTGGCAGACAGACTTATGCTGGAACCTGGG 1201

QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGl 318
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1202 TTTTAGCAAGTTTTCAGATAGAACACCTTTGACTTTATGGAGAAATATTCACCTGGA 1261

QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTrpGlnArgPheAlaValMetAl 338
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1262 AGGAAGACTAACTCTTTTCAGAGAGAGTAGGCGAGTATCAGAGGATGGAGTGTGTC 1321

QY 338 aGluThrTrpAspAsnValPheThrLeuAspAlaAspPhe 351
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1322 AAGTCCAAACAGAAATCTTTTACCTTGGATGCTGACTTC 1361
```

RESULT 6

```
US-09-962-665-9
; Sequence 9, Application US/09962665
; Patent No. 6537759
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: POLYPOLYGLUTAMATE SYNTHETASE GENE SEQUENCE
; TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING THE
; TITLE OF INVENTION: TREATMENT OF DISEASE
; FILE REFERENCE: 11926-015004
; CURRENT APPLICATION NUMBER: US/09/962,665
; CURRENT FILING DATE: 2001-09-24
```

```
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
```

```
; LENGTH: 2500
```

```
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
```

```
; FEATURE:
```

```
; NAME/KEY: misc_feature
```

```
; LOCATION: 128..1464
```

```
; OTHER INFORMATION: n = g or a
```

```
; NAME/KEY: misc_feature
```

```
; LOCATION: 189..
```

```
; OTHER INFORMATION: n = t or g
```

```
; NAME/KEY: misc_feature
```

```
; LOCATION: 524
```

```
; OTHER INFORMATION: n = c or g
```

```
; NAME/KEY: misc_feature
```

```
; LOCATION: 1399..
```

```
; OTHER INFORMATION: n = t or a
```

```
; NAME/KEY: misc_feature
```

```
; LOCATION: 1636..1738, 2259
```

```
; OTHER INFORMATION: n = c or t
```

```
US-09-962-665-9
```

```
Alignment Scores:
```

```
Pred. No.: 2,09e-179 Length: 2500
Score: 1460.50 Matches: 277
Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
DB: 4 Gaps: 1
```

```
US-10-698-228-1 (1-351) x US-09-962-665-9 (1-2500)
```

```
QY 4 ProGluArgProGluAlaAlaGlyLeuAspGluArgSerSerSerAspThrAs 23
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 302 CCTGAGCGGAGACCGCGTCTGGCCAGCAAGCCGAGGAGGATCTTCCAGAGCCAC 361

QY 23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 362 GGAGCCGAAACTAAAGCAGCTGCCCGCGTGGAGGATGAGCGCTGCTGAGAGAAA 421

QY 38 rSerArgArgPheValIlePheProIleGlnTrpProAspIleTrpLysMetTyIysGl 58
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 422 CCCCCTGGCTTGTCTATCTTCCCATCAGTACCATGATATCTGGCAGATGTATAAGAA 481

QY 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 482 GGCAGAGGCTTCTCTTTTGGACCGCGAGGAGTTGACCTCTCNAAGGACATTCAGCAGCTG 541

QY 78 pAsnLysLeuLysAlaAspGluLysTrpPheLeuSerHisIleLeuAlaPhePheAlaAl 98
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 542 GGAATCCCTGAAACCCGAGAGAGATATTTATATCCATGTTCTGGCTTCTTTGAGC 601

QY 98 aSerAspGlyLeuValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 602 AAGCATGGATAGTAATGAAACTTGGTGGAGGATTTAGCCAGAAAGTTTCAGATTAC 661

QY 118 oGluAlaArgCysPheTrpGlyPheGlnIleLeuLeuValHisSerGluMetTy 138
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 662 AGAAGCCCGCTGTTTCTATGGCTTCCAAATGGCATGGAACATACATATTCGAAATGTA 721

QY 138 rSerLeuLeuLeuAspThrTrpIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
Db 722 TAGCTCTTATTGACACTTACATAAAGATCCCAAGAAAGGAATTTCTTCAATGC 781
Qy 158 alleGluThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
Db 782 CATTTGAAACGATCGCTTGTGTCAAGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 841
Qy 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePh 198
Db 842 CAAAGAGGCTACCTATGTTGTAACGCTGTGTAGGCTTTGCTGCAAGGAGCAATTTCTT 901
Qy 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
Db 902 TTCGGGTTCTTTTGGTGTGATATCTGGCTCAAGAAAGGAGGAGGAGGAGGAGGAGGAG 961
Qy 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
Db 962 ATTTTCTAATGAACCTTATTAGCAGAGATGAGGCTTTACACTGTGATTTTCTGCTGAT 1021
Qy 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
Db 1022 GTTCAACACCTGGTACACAAACCATCGGAGGAGAGAGTAAGAGAAATAATTATCAATGC 1081
Qy 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
Db 1082 TGTTCCGATAGACAGAGTTCTCTACTGAGGCTTGCCTGTGAAGCTCAATGGATGAA 1141
Qy 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAlaAspArgLeuLeuValGluLeuG 298
Db 1142 TTGCACCTAATGAACCAATACATTGAGTTTGTGGCAGACAGACTTATGCTGGAACCTGG 1201
Qy 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG 318
Db 1202 TTTTAGCAAGGTTTTCAGAGTACAGAACCCATTTTGACTTTATGAGAGAAATATTTCACTGGA 1261
Qy 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
Db 1262 AGGAAAGACTAACTTTTGTGAGAGAGAGTAGGCGAGATACAGAGGATGGAGTGATGTC 1321
Qy 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1322 AAGTCCACACAGAAATCTTTTACCTTGGATGCTGACTTC 1361
```

RESULT 7

```
US-09-023-655-1370
; Sequence 1370, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREMITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
```

```
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1370:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2500 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: G36154
; US-09-023-655-1370

Alignment Scores:      2.09e-179      Length:      2500
Pred. No.:            1460.50          Matches:     277
Score:                88.42%           Conservative: 36
Percent Similarity:   78.25%           Mismatches:  35
Best Local Similarity: 80.20%           Indels:       6
Query Match:          4                Gaps:         1
DB:
```

```
US-10-698-228-1 (1-351) x US-09-023-655-1370 (1-2500)

Qy 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
Db 302 CCTGAGCGGACCGCGCTCTGCGCAGAACACCGCGAGGAGGATCTTCCAGGAGCCAC 361
Qy 23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
Db 362 GGAGCGGAACCTAAAGCAGCTGCCCCGGCGTGGAGGATGAGCGCTGTGAGAGAAA 421
Qy 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysG 58
Db 422 CCCCCGCGCTTGTGTCATCTTCCCATCGATACATCATATCTGGCAGATGTATAAGAA 481
Qy 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
Db 482 GGCAGAGGCTTCTTTTGGACCGCGGAGGAGTGTACCTTCCCAAGGACATTCAGCATG 541
Qy 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAl 98
Db 542 GGAATCCCTGAACCCGAGGAGAGATATTTATATCCATGTTCTGGCTTTCTTTGACG 601
Qy 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
Db 602 AAGCATGGCATAGTAATAAGAACTTGGTGGAGCGATTAGCCAAGAAAGTTCAAGATTAC 661
Qy 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
Db 662 AGAAGCCCGCTGTTCTTATGCTTCCAAATTTGCCAAACCAACATACATCTCGAAATGA 721
Qy 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
Db 722 TAGTCTTCTTATTGACACTTACATAAAGATCCCAAGAAAGGGAATTTCTTCAATGC 781
Qy 158 aIleGluThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
Db 782 CATTTGAAACGATCGCTTGTGTCAAGAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 841
Qy 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePh 198
Db 842 CAAAGAGGCTACCTATGTTGTAACGCTGTGTAGGCTTTGCTGCAAGGAGCAATTTCTT 901
Qy 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
Db 902 TTCGGGTTCTTTTGGTGTGATATTTCTGCTCAAGAAACGAGGAGTGTATGCTGCCTCAC 961
```

Qy 218 rPheSerAsnGluLeuIleSerArgAspGluGluValLeuHisCysAspPheAlaCysLeuMe 238
Db 962 ATTTTCTTAATGAACCTTATTAGCAGAGATGAGGGTTTACACTGTGATTTTCTGCTCGCTGAT 1021
Qy 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluLeuIleValAspAl 258
Db 1022 GTTCAACACCTGGTACACAAACCATCGAGGAGAGAGTAAAGANAATATTATCAATGC 1081
Qy 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
Db 1082 TGTTCCGATAGAACAGAGTTCTCTCACTAGAGCCTTGCCTGTGAAGCTCATTTGGATGA 1141
Qy 278 nCysIleLeuMeLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGl 298
Db 1142 TTGCACCTTAATGAACATACATTTGAGTTTGTGGCAGACAGACTTATGCTGGAACCTGG 1201
Qy 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGl 318
Db 1202 TTTTAGCAAGGTTTTCAGAGTAGAGAACCCATTGTGACTTTATGGAGAAATATTTCACCTGA 1261
Qy 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
Db 1262 AGGAAGACTAATCTTTTCAGAGAGAGAGTAGGCGAGTATCAGAGGATGGAGTGTGATGC 1321
Qy 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1322 AAGTCCAAACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1361

RESULT 8

US-09-963-333-9
; Sequence 9, Application US/09963333
; Patent No. 6664062
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: THYMIDINE SYNTHASE GENE SEQUENCE VARIANCES
; TITLE OF INVENTION: HAVING UTILITY IN DETERMINING THE TREATMENT
; FILE REFERENCE: 11926-015002
; CURRENT APPLICATION NUMBER: US/09/963,333
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 128, 1464
; OTHER INFORMATION: n = g or a
; NAME/KEY: misc_feature
; LOCATION: 189
; OTHER INFORMATION: n = t or g
; NAME/KEY: misc_feature
; LOCATION: 524
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc_feature
; LOCATION: 1399
; OTHER INFORMATION: n = t or a
; NAME/KEY: misc_feature
; LOCATION: 1636, 1738, 2259
; OTHER INFORMATION: n = c or t
; US-09-963-333-9

Alignment Scores: 2.09e-179 Length: 2500
Pred. No.: 1460.50 Matches: 277
Score: 1460.50
Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
DB: Gaps: 1
US-10-698-228-1 (1-351) x US-09-963-333-9 (1-2500)
Qy 4 ProGluArgProGluAlaAlaGlyLeuAspGlu-ArgSerSerSerAspThrAs 23
Db 302 CCTGAGCGGACCCCGCTCTGCGCAGCAAGACCCGAGGAGGATCTTCCAGAGGCCAC 361
Qy 23 nGluSerGluLeuLysSer-----AsnGluGluProLeuLeuArgLysSe 38
Db 362 GGACCCGMAAATAAGCAGCTGCCCGCGCTGGAGGATGAGCGCTCTGTAGAGAAA 421
Qy 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGl 58
Db 422 CCCC CGCGCTTTGTTCATCTTCCCATCGAGTACCATGATATCTGCAGATGTATAAGAA 481
Qy 58 nAlaGlnAlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
Db 482 GGCAGAGGCTTCTTTTGGACCGCGAGGAGGTTGACCTCTCNAAGGACATTCAGCACCTG 541
Qy 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAl 98
Db 542 GGAATCCCTGAAACCCGAGGAGAGATTTTATATCCCATGTTCTGGCTTCTTTTCGAGC 601
Qy 98 aSerAspGlyIleValAsnGluValGluArgPheSerGlnGluValGlnValPr 118
Db 602 AAGCGATGGCATAGTAAATGAATGAACTTGTGGAGCGATTTAGCCAAAGAGTTTCAGATTAC 661
Qy 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
Db 662 AGAAGCCGCTGTTTCTATGCTTCCAAATTTGCCATGCAAAACATACATCTCTGAATGTA 721
Qy 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
Db 722 TAGTCTTCTTATTGACACTTACATAAAGATCCCAAGAAAGAGGAAATTTCTCTTCAATGC 781
Qy 158 aIleGluThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
Db 782 CATTTGAACGATGCTTGTGTCAAGAGAGGACACTGGGCTTGGCGCTGGATGGGGA 841
Qy 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePh 198
Db 842 CAAAGAGGCTACCTATGTTGAACCGTGTGTAGCCCTTGTGTCAGTGGAGGCAATTTCTT 901
Qy 198 eSerGlySerPheAlaAlaIlePheTyrPheLysLysArgGlyLeuMetProGlyLeuTh 218
Db 902 TTCCGGTTCTTTTGGCTGATATTTCTGCTCAAGAAACGAGGACTGATCCCTGGCCTCAC 961
Qy 218 rPheSerAsnGluLeuIleSerArgAspGluGluLeuHisCysAspPheAlaCysLeuMe 238
Db 962 ATTTTCTTAATGAACCTTATTAGCAGAGATGAGGGTTTACACTGTGATTTTCTGCTCGCTGAT 1021
Qy 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
Db 1022 GTTCAACACCTGGTACACAAACCATCGAGGAGAGAGTAAAGANAATATTATCAATGC 1081
Qy 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
Db 1082 TGTTCCGATAGAACAGAGTTCTCTCACTGAGGCTTGGCTGTGGAAGCTCATTTGGGATGAA 1141
Qy 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuGl 298
Db 1142 TTGCACCTTAATGAACATACATTTGAGTTTGTGGCAGACAGACTTATGCTGGAACCTGG 1201
Qy 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGl 318

Db 1202 TTTTACAGGTTTTCAGAGTAGAGAACCAATTTGACTTTATGGAGATATTTCACTGGA 1261
 Qy 318 uGlyLysThrAsnPhaPheGluLysArgValSerGluTyrGlnArgPheAlaValMetal 338
 Db 1262 AGGAAAGACTAACTTTTTCAGAGAGAGTAGGCGAGATTCAGAGGATGGAGTGATGTC 1321
 Qy 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
 Db 1322 AAGTCCACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1361

RESULT 9

US-09-962-677-9
 ; Sequence 9, Application US/09962677
 ; Patent No. 6759200
 ; GENERAL INFORMATION:
 ; APPLICANT: Stanton, Jr., Vincent P.
 ; TITLE OF INVENTION: THYMIDINE PHOSPHORYLASE GENE SEQUENCE
 ; TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING
 ; TITLE OF INVENTION: THE TREATMENT OF DISEASE
 ; FILE REFERENCE: 11926-015003
 ; CURRENT APPLICATION NUMBER: US/09/962,677
 ; CURRENT FILING DATE: 2001-09-24
 ; PRIOR APPLICATION NUMBER: 09/658,659
 ; PRIOR FILING DATE: 2000-09-08
 ; PRIOR APPLICATION NUMBER: 09/596,033
 ; PRIOR FILING DATE: 2000-06-15
 ; PRIOR APPLICATION NUMBER: 09/357,743
 ; PRIOR FILING DATE: 1999-07-20
 ; PRIOR APPLICATION NUMBER: 09/357,024
 ; PRIOR FILING DATE: 1999-07-19
 ; PRIOR APPLICATION NUMBER: 60/093,484
 ; PRIOR FILING DATE: 1998-07-20
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 9
 ; TYPE: DNA
 ; LENGTH: 2500

ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: 128..1464
 OTHER INFORMATION: n = g or a
 NAME/KEY: misc feature
 LOCATION: 189
 OTHER INFORMATION: n = t or g
 NAME/KEY: misc feature
 LOCATION: 524
 OTHER INFORMATION: n = c or g
 NAME/KEY: misc feature
 LOCATION: 1399
 OTHER INFORMATION: n = t or a
 NAME/KEY: misc feature
 LOCATION: 1636..1738, 2259
 OTHER INFORMATION: n = c or t
 US-09-962-677-9

Alignment Scores:

Pred. No.: 2,09e-179 Length: 2500
 Score: 1460.50 Matches: 277
 Percent Similarity: 88.42% Conservative: 36
 Best Local Similarity: 78.25% Mismatches: 35
 Query Match: 80.20% Indels: 6
 DB: 4 Gaps: 1

US-10-698-228-1 (1-351) x US-09-962-677-9 (1-2500)

Qy 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerAspThrAs 23
 Db 302 CTTGAGCGGACCCCGGTCTTGGCCAGACGACCGGAGGAGGATCTTCCAGGAGCCAC 361
 Qy 23 nGluSerGluLeuLysSer-----AsnGluProLeuLeuArgLysSe 38
 Db 362 GGAGCGGAAAACTAAAGCAGCTGCCCGCGGTGGAGGATGAGCCGCTGCTGAGAGAAA 421

Qy 38 rSerArgArgPheValIlePhePheProIleGlnTyrProAspIleTrpLysMetTyrLysG 58
 Db 422 CCCC CGCGCGTGTGTATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481
 Qy 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
 Db 482 GGCAGAGGCTCTCTTTTGGACCGCGGAGGAGGTGACCTCTCNAGAGCATCTCAGCACTG 541
 Qy 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaA 98
 Db 542 GGAATCCCTGAAACCCGAGGAGAGATATTTATATCCCATGTTCTGGCTTTCTTTCAGC 601
 Qy 98 aSerAspGlyLylleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
 Db 602 AAGCGATGGCATAGTAATAAGAACTTGTGTGGAGCGATTTAGCCCAAGAAAGTTTCAGATTAC 661
 Qy 118 oGluAlaAatqCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
 Db 662 AGAAGCCCGCTGTTCATGGCTTCCAAATTTGCATGGAACATACATCTGAAATGTA 721
 Qy 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
 Db 722 TAGTCTCTTATTGACACTTACATAAAGATCCCAAGAAAGGAAATTTCTTCAATGC 781
 Qy 158 alieGluThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
 Db 782 CATTGAAACGATGCTTGTCTCAAGAAAGAGGAGAGCTGGGCTTGGCTGGATGGGA 841
 Qy 178 pArgLysSerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePh 198
 Db 842 CAAAGAGGCTACCTATGTTGCAACGTGTGTAGCCCTTGTGTCAGTGAAGGCAATTTCTT 901
 Qy 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
 Db 902 TTCGGTCTTTTTCGTCGATATTCGCTCAAGAAACGAGGAGTGTGCTTGGCTTGCCTCAC 961
 Qy 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
 Db 962 ATTTTCTTAATGAACCTTATTAGCAGAGATGAGGGTTTACACTGTGATTTTGTCTGCTGAT 1021
 Qy 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
 Db 1022 GTTCAAAACACCTGTTACAAACCATCGAGGAGAGTAGAGAGAAATAATTATCAATGC 1081
 Qy 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
 Db 1082 TGTTGGATAGAACAGGAGTTCTCACTGAGGCCCTTGCCTGTGAAGCTCATTTGGATGAA 1141
 Qy 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGl 298
 Db 1142 TTGCACCTTAATGAAGCAATACATTGAGTTTGTGGCAGACAGACTTATGCTGGAACCTGG 1201
 Qy 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGl 318
 Db 1202 TTTTAGCAGGTTTTCAGAGTAGAGAACCCATTTGACTTTTATGAGAGATATTTTCACTGGA 1261
 Qy 318 uGlyLysThrAsnPhaPheGluLysArgValSerGluTyrGlnArgPheAlaValMetal 338
 Db 1262 AGGAAAGACTAACTTTTTCAGAGAGAGTAGGCGGAGTATCAGAGGATGGAGGTGATGTC 1321
 Qy 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
 Db 1322 AAGTCCACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1361

RESULT 10

US-09-949-016-145
 ; Sequence 145, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

```
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 145
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-145

Alignment Scores:
Pred. No.: 2,09e-179 Length: 2500
Score: 1460.50 Matches: 277
Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
DB: 4 Gaps: 1

US-10-698-228-1 (1-351) x US-09-949-016-145 (1-2500)

Qy 4 ProGluArgProGluAlaAGlyLeuAspGluArgSerSerSerAspThrAs 23
Db 302 CCGAGCGGACCGCGGCTCTGCGCAGCAAGACGCGGAGGAGATCTCCAGGAGCCAC 361
Qy 23 nGluSerGluLeuLysSer-----AsnGluGluProLeuLeuArgLysSe 38
Db 362 GAGCGGAAACTAAGCACTGCCCCCGCGCGAGAGTAGCGGCTGCTGAGAGAAA 421
Qy 38 rSerArgArgPheValIlePheProLeuGlnTyrProAspIleTrpLysMetTyrLysG 58
Db 422 CCCCAGCGGCTTTGTCATCTTCCCATCGAGTACCATGATATCTGCGAGATGTATAAGAA 481
Qy 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
Db 482 GCGAGAGGCTTCTTTTGGACCGCGGAGGAGGTGACCTCTCCAAGGACATTCAGCACTG 541
Qy 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAl 98
Db 542 GGAATCCCTGAAACCGGAGAGAGATATTTATATCCATGTTCTGCGCTTCTTTCAGC 601
Qy 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
Db 602 AAGCGATGGCATAGTAATAATGAAACTTTGGTGGAGCGATTAGCCAAAGAGTTCAGATTAC 661
Qy 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
Db 662 AGAAGCGCGCTGTTCTATGGCTTCCAAATTTGCCATGCAAAACATACATTCGAAATGTA 721
Qy 138 rSerLeuLeuLeuAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
Db 722 TAGTCTTCTTATTGACACTTACATAAAGATPCCAAAGAAAGGGAATTTCTTCAATGC 781
Qy 158 aIleGluThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
Db 782 CATTGAACCATGCTTGTGTCAAGAAAGAGGAGACTGGGCTTGGCTGGATGGGA 841
Qy 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePh 198
Db 842 CAAAGAGGCTACCTATGCTGAACGCTGTGTAGCTTTGCTGCTGCAAGGCAATTTCTT 901
Qy 198 eSerGlySerPheAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
Db 902 TTCGGGTCTTTTTCGCTCGCATATTTCTGCTCAAGAAACGAGGAGTGTGCTGCGCTCAC 961
Qy 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
```

```
Db 962 ATTTTCTAATGAACCTTATTAGCAGAGATGAGGGTTTACACTGTGATTTTGTGCTGCTGAT 1021
Qy 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
Db 1022 GTTCAAAACACCTGCTGACAAACCATCGAGGAGAGAGTAAGAGAAATAATTAATCAATGC 1081
Qy 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
Db 1082 TGTTCCGATGAACACAGGAGTTCTCTCAGGAGGCTTGGCTGTGAAGCTCATTTGGGATGAA 1141
Qy 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuG 298
Db 1142 TTGCATCTTAATGAAGCAATACATTAAGTTGTGCGACAGACATTAATGCTGGAATGGG 1201
Qy 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG 318
Db 1202 TTTTAGCAAGGTTTTCAGAGTAGAGAACCCATTTGACTTTATGGAGAAATATTTCACTGGA 1261
Qy 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetal 338
Db 1262 AGGAAAGACTTAATCTTTTGAGAGAGAGTAGGCGAGTATCAGAGGATGGGAGTGATGTC 1321
Qy 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1322 AAGTCCACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1361

RESULT 11
US-08-307-499-1/c
; Sequence 1, Application US/08307499
; Patent No. 5651972
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: U.S.A.
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/307,499
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,241
; FILING DATE: 1-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,630
; FILING DATE: 29-JUN-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/342,212
; FILING DATE: 21-APR-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF35.1.FWCC1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
```

LENGTH: 14176 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: unknown
 MOLECULE TYPE: DNA (genomic)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 3852..4226
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 4585..4887
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 5131..5310
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 5760..5912
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 6786..7130
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 10148..10513
 US-08-307-499-1

Alignment Scores:
 Pred. No.: 1.37e-145 Length: 14176
 Score: 1210.00 Matches: 233
 Percent Similarity: 82.01% Conservative: 36
 Best Local Similarity: 71.04% Mismatches: 58
 Query Match: 66.45% Indels: 2
 DB: 1 Gaps: 1

US-10-698-228-1 (1-351) x US-08-307-499-1 (1-14176)

QY 24 GluSerGluLeuLysSerAsnGluGluProLeuLeuArgLysSerSerArgPheVal 43
 DB 10842 AAAATGATTTTATACAGAAATGGAGCCATTTCTTCAAGAGTCTGATTCAGTTCGTT 10783
 QY 44 IlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGlnAlaSerPhe 63
 DB 10782 ATTTCCCTTAAGTATCATGATATCTGGAATGATTAACATCATGAGTGGCAGTTT 10723
 QY 64 TrpThrAlaGluGluValAlaSerLysLeuSerLysLeuProHisTrpAsnLysLeuAla 83
 DB 10722 TGGACCGTTGACAGAGTAGATTATCAAAAGATTAGATGATGGGATAAATTAACATA 10663
 QY 84 AspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerArgGlyIleVal 103
 DB 10662 GACGAAAAATACTTTATAAACATATACTAGCATTTTTCATCTAGTGTGATTTGTA 10603
 QY 104 AsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAlaArgCysPhe 123
 DB 10602 AATGAGAAATTTACGGAAGATTTTATGTGGATGACAGTGTTCAGAGCCAGCATGTTTC 10543
 QY 124 TyrGlyPheGlnIleLeuGluAsnValHisSerGluMetTyrSerLeuLeuLeuAsp 143
 DB 10542 TATGGATTTCAATAGTATGGAATATTTTCAATTCAGAAATGATATAGTTTATTAATAGAT 10483
 QY 144 ThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGluThrMetPro 163
 DB 10482 ACATATGTAAGATAATAAGAAAAATGCAATTTATTAACGCTATAGAAACAATGGAA 10423
 QY 164 TyrValLysLysLysAlaAspTyrAlaLeuArgTyrIleAlaAspArgLysSerThrPhe 183
 DB 10422 TCGTAAAAAGAAAGCTGATTCGGCCAGAAAATGGATATCTAGCAACAAG---GTATAT 10366
 QY 184 GlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGlySerPheAla 203
 DB 10365 GGNAAAGAGTAGTACATTTGAGCTGTGAGGGAATATCTTTCTGTTCAATTTGCT 10306
 QY 204 AlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeu 223

DB 10305 GCTATATTTTGGATAAAAAACAGAGATTGATCCCGGATTAAACATTTTCTAATGAACATA 10246
 QY 224 IleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGlnTyrLeuVal 243
 DB 10245 ATAAGTAGAGACGAAGGTTTACATTTGCTGTTTAAATGTTTAAACATTTATTA 10186
 QY 244 AsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLysIleGluGln 263
 DB 10185 CATCCACCATCTAAGGAAGTTTATAACGTCGATAATCATGTATGCGGTTAATATAGAAAAG 10126
 QY 264 GluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIleLeuMetLys 283
 DB 10125 GAGTTTTTGACAGTGTCTATTCGGTGCATCTTATAGGTATGAATTTGTTTAAATGCT 10065
 QY 284 GlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSerLysValPhe 303
 DB 10065 CAGTATATAGAAATTCGTCGAGATAGATTATTAACAGAGTTAGGTTGTGAAAAAGT--CTC 10008
 QY 304 GlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLysThrAsnPhe 323
 DB 10007 AATGTATATATCTCTTTTAGCTTTTATGGAGTATATATCATCAAGAGGTAAGACTAATTT 9948
 QY 324 PheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThrThrAspAsn 343
 DB 9947 TTCGAACACGAGTTAGTATATCAAAAGATGGGGTGTGTACAAATAAAGAGAGAT 9888
 QY 344 ValPheThrLeuAspAlaAspPhe 351
 DB 9887 ATAATTTCTACGGATATAGATTTT 9864

RESULT 12
 US-08-307-499-14
 ; Sequence 14, Application US/08307499
 ; Patent No. 5651972
 ; GENERAL INFORMATION:
 ; APPLICANT: Moyer, Richard W.
 ; APPLICANT: Vi uela, Eladio
 ; APPLICANT: Gibbs, E.P.J.
 ; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
 ; TITLE OF INVENTION: Live Vaccine Vector
 ; NUMBER OF SEQUENCES: 60
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: David R. Saliwanchik
 ; STREET: 2421 N.W. 41st Street, Suite A-1
 ; CITY: Gainesville
 ; STATE: Florida
 ; COUNTRY: U.S.A.
 ; ZIP: 32606
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/307,499
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/908,241
 ; FILING DATE: 1-JUL-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/908,630
 ; FILING DATE: 29-JUN-1992
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/342,212
 ; FILING DATE: 21-APR-1992
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Saliwanchik, David R.
 ; REGISTRATION NUMBER: 31,794
 ; REFERENCE/DOCKET NUMBER: UP35.1.FWCCI
 ; TELECOMMUNICATION INFORMATION:


```

; Patent No. 6217882
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; TITLE OF INVENTION: Live Vaccine Vector
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: U.S.A.
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/299,268
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/901,127
; FILING DATE:
; APPLICATION NUMBER: US 07/908,241
; FILING DATE: 1-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,630
; FILING DATE: 29-JUN-1992
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/342,212
; FILING DATE: 21-APR-1992
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF35.1.FWCCI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14176 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3852..4226
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 4585..4887
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5131..5310
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5760..5912
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 6786..7130
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 10148..10513
; US-09-299-268-1
; Alignment Scores:
; Pred. No.: 1.37e-145 Length: 14176

```



```

QY 144 ThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGluThrMetPro 163
DB 3695 ACATATGTAAAGATAATATAGAAAAATGCAATTTATTTAAACGCTATAGAAACAATGGAA 3754
QY 164 TyrValLysLysLysAlaAspTrpAlaLeuAuaGTrpIleAlaAspArgLysSerThrPhe 183
DB 3755 TCGGTAAATAAGAAAGCTGATTGGGCCAGAAATGGATATCTAGCAACAG---GTATAT 3811
QY 184 GlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGlySerPheAla 203
DB 3812 CGAGAAAGAGTAGTACCAATTCAGCTGTGGAGGAATATCTTTCTGGTTCAATTTGCT 3871
QY 204 AlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeu 223
DB 3872 GCTATATTTGGTAAATAAAAAAGAGGATTCATGCCCGGATTAACATTTCTAATGAAC 3931
QY 224 IleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGlnTyrLeuVal 243
DB 3932 ATAAGTAGAGAGCAAGGTTTACATGTGATTTTGCCTGTTTAAATGTTTAAACATTTATTA 3991
QY 244 AsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLysIleGluGln 263
DB 3992 CATCCACCATCTAAGGAAGTTATACCTCGATAATCATTTGATGCGGTTAATATAGAAAG 4051
QY 264 GluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIleLeuMetLys 283
DB 4052 GAGTTTTCACAGTTGCTATTCCTCGGTGGATCTTATAGGTATGAATTTGTTTAAATGCT 4111
QY 284 GlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSerLysValPhe 303
DB 4112 CAGTATATAGAATTCGTCCAGATAGATTATTAAACAGAGTTAGTTGTGAAAGT--CTC 4169
QY 304 GlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLysThrAsnPhe 323
DB 4170 AATGTATATAATCCCTTTAGCTTTATGGAGTATATATCTAGAGGTAAAGTAAATTTT 4229
QY 324 PheGluLysArgValSerGlnTyrGlnArgPheAlaValMetAlaGluThrThrAspAsn 343
DB 4230 TTCCGAACGAGGTTAGTCAATATCAAAAGATGGGGGTGTTTACAAATAAAGAGAGAT 4289
QY 344 ValPheThrLeuAspAlaAspPhe 351
DB 4290 ATATTTTCTACGATATAGATTTT 4313

RESULT 15
US-09-248-796A-3914
; Sequence 3914, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE OF INVENTION: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 3914
; LENGTH: 1236
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-3914

Alignment Scores:
Pred. No.: 5,45e-137 Length: 1236
Score: 1131.00 Matches: 210
Percent Similarity: 78.21% Conservative: 52
Best Local Similarity: 62.69% Mismatches: 67
Query Match: 62.11% Indels: 6

```

```

DB: 4 Gaps: 3
US-10-698-228-1 (1-351) x US-09-248-796A-3914 (1-1236)
QY 23 AsnGluSerGluIleLysSerAsnGlu-----GluProLeuLeuArgLysSerSerArg 40
DB 229 AACAGACACAAAGTTAAACCAAGAGAGTCCACGAGCAATGGCTTACACACAGACAGACG 288
QY 41 ArgPheValIlePhePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
DB 289 CGTCATACCATATATCCGATAAAGTATCCCGAGTATGGCAGTCTCTATAAAAGTCGTGA 348
QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrp---Asn 79
DB 349 GCCTCATCTCGACTGCTGAAGAGCTTGACTGTAGTAAAGATTGGAGTATGGAATATAT 408
QY 80 LysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAsp 99
DB 409 AAGATGAACGAAACGAAAGATTTCTTTATCTCTCGAGTTTGGCATCTTTTGTGTCATCG 468
QY 100 AspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGlu 119
DB 469 GATGGTATTGTTAATGAGAACTTGGTGGAAAATTTCTGTGCAGAGTGCAAAATACCCGAG 528
QY 120 AlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSer 139
DB 529 GCTAAACTGCTTTTACAAAGTTTCAGATCATGATGAGAAACATTCATTTCAGAGACTTATTCG 588
QY 140 LeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIle 159
DB 589 TTGTTAATAGAAACATATTTCAAGACCCCGAGACAGACACTTTTGTGTTAATGCTATT 648
QY 160 GluThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArg 179
DB 649 GACAAATCCCATTTATAAGAAAAGCCGATTTGGCAATCAGATGCATCCAGACGAG 708
QY 180 LysSerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSer 199
DB 709 GATGCTTTGTATGCTGAGAGATTTGTTGCTATTTGCTGCTGTTGAAGGTATATTTTTCAGT 768
QY 200 GlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPhe 219
DB 769 GGCTCGTTTGCAGCAATTTTCTGTTTAAAGAAACGGGGTTTGATGCCGGGGTTGACATT 828
QY 220 SerAsnGluLeuIleSerArgAspGluGlyLysHisCysAspPheAlaCysLeuMetPhe 239
DB 829 TCCATATGAATTAATTTGTCAGACGAGAGGATTCACACAGATATATGCTGTTTATTTT 888
QY 240 GlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaVal 259
DB 889 TCATATTTAAAAATAAACCTTCTTCAGAAATAATTCAGAAATAATCATCCCGAGGCTGTT 948
QY 260 LysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCys 279
DB 949 GATATCGAGAAAGATATTTTTCAGCGATCGCTTCCAGTATCTCTTATTTGGGAATGAATGT 1008
QY 280 IleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPhe 299
DB 1009 GACTTGTATGCCAGTATGTGAATTTGTTGCTGATAGATTACTAGTAGCTTTTGGAAAT 1068
QY 300 SerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGly 319
DB 1069 AAAAAGTACTACAATGTTACCAACCCCATTTTATGTTATGAAATAATTTCTCTGCTGCT 1128
QY 320 LysThrAsnPhePheGluLysArgValSerGlnTyrGlnArgPheAlaValMetAlaGlu 339
DB 1129 AAAACTAATTTCTTTGAAAAGAGAGTCTCTGATTATCAAAAGGCTGTTGTTGATGAGAAA 1188
QY 340 Thr-----ThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB 1189 GTGGAATAAAGAAATGAGAAAACCTGGATTTGTTGATCAAGATTTTC 1233

```

Search completed: June 16, 2005, 18:40:38
Job time : 244 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 16, 2005, 13:06:58 ; Search time 703 Seconds
(without alignments)
3099.358 Million cell updates/sec

Title: US-10-698-228-1

Perfect score: 1821

Sequence: 1 MGDPRPEAAGLDQDERSSS.....QRFVMAETTDNVFTLDADF 351

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 6054689 seqs, 3103772919 residues

Total number of hits satisfying chosen parameters: 12109378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n_model -DRV=xlh
-Q=/cgn2_1/USPTO.spool/US10698228/runat_14062005_161418_28424/app.query.fasta_1.519
-DB=Published Applications NA -QWTF=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPTXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10698228 @CGN 1 1 725 @runat_14062005_161418_28424
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/ptodata/2/pubpna/US10J_NEW_PUB.seq:*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1821	100.0	1053	18	US-10-698-228-2
2	1821	100.0	1053	18	Sequence 2, Appli
3	1821	100.0	1081	18	Sequence 12, Appl
4	1821	100.0	4955	17	US-10-698-228-4
5	1821	100.0	4955	17	Sequence 4, Appli
6	1821	100.0	4955	17	Sequence 71, Appl
7	1460.5	80.2	1989	9	US-10-172-118-71
8	1460.5	80.2	2216	15	US-10-342-887-71
9	1460.5	80.2	2500	9	US-10-698-228-3
10	1460.5	80.2	2500	9	Sequence 3, Appli
11	1460.5	80.2	2500	9	Sequence 505, App
12	1460.5	80.2	2500	20	Sequence 342, App
13	1460.5	80.2	2500	21	Sequence 724, App
14	1460.5	80.2	2500	9	Sequence 1169, Ap
15	1460.5	80.2	2500	9	Sequence 1827, Ap
16	1460.5	80.2	2500	9	Sequence 1370, Ap
17	1443	79.2	2113	21	Sequence 458, App
18	1425.5	78.3	1328	18	Sequence 3751, Ap
19	1280	70.3	186854	21	Sequence 4196, Ap
20	1230	67.5	1218	15	Sequence 4854, Ap
21	1220.5	67.0	1292	17	Sequence 75, Appl
22	1192.5	65.5	1146	15	Sequence 34, Appl
23	1192	65.5	1242	16	Sequence 7245, Ap
24	1176.5	64.6	963	17	Sequence 27809, A
25	1168.5	64.2	1200	17	Sequence 2245, Ap
26	1161.5	63.8	1173	17	Sequence 6537, Ap
27	1161.5	63.8	1206	17	Sequence 34052, A
28	1151	63.2	1314	15	Sequence 45857, A
29	1151	63.2	1314	15	Sequence 46252, A
30	1151	63.2	3314	15	Sequence 2145, Ap
31	1151	63.2	3450	15	Sequence 1245, Ap
32	1147.5	63.0	1586	18	Sequence 6245, Ap
33	1133.5	62.2	1169	18	Sequence 245, App
34	1133.5	62.2	2992	20	Sequence 5245, Ap
35	1132.5	62.2	1725	19	Sequence 73024, A
36	1128.5	62.0	1255	19	Sequence 31361, A
37	1125.5	61.8	1791	20	Sequence 11182, A
38	1114.5	61.2	1248	18	Sequence 11709, A
39	1113	61.1	1146	9	Sequence 63321, A
40	1097.5	60.3	1361	18	Sequence 128843,
41	1092.5	60.0	1569	20	Sequence 174, App
42	1062	58.3	957	20	Sequence 40471, A
43	1043	57.3	1306	20	Sequence 15462, A
44	1036.5	56.9	1420	17	Sequence 43, Appl
45	1036.5	56.9	3420	17	Sequence 1378, Ap
					Sequence 1145, Ap
					Sequence 145, App

ALIGNMENTS

RESULT 1

US-10-698-228-2
; Sequence 2, Application US/10698228
; Publication No. US20040072253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06

```
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-2

Alignment Scores:
Pred. No.:          9.6e-217      Length:      1053
Score:              1821.00      Matches:     351
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%  Mismatches:  0
Query Match:        100.00%      Indels:      0
DB:                  18          Gaps:         0

US-10-698-228-1 (1-351) x US-10-698-228-2 (1-1053)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
DB 1 ATGGGCGACCCGGAAAGCGGAGCGCGCGCTGGATCAGATGAGATCATCTTCA 60
QY 21 AspThrAsnGluSerGluLeuLysSerAsnGluGluProLeuLeuArgLysSerArg 40
DB 61 GACACCAACGAAAGTGAATGAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
DB 121 CGGTTTGTCTATCTTCCATCCAGTACCTCCCTGATATTGGAAATGATATAACAGGCACAG 180
QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
DB 181 GCTTCCTTCTGCACAGCAGAGAGGTGCGACTTATCAAGAGATCTCCTCACTGGAACAAG 240
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
DB 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCTCACATCTTAGCCTTTTTCAGGCAGGTAT 300
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 360
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 361 CGCTGTCTTCTAGGCTTTCAAAATTCATCGAATGTTCACTCAGAGATGTACAGTTTG 420
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTATGCAATTGAA 480
QY 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
DB 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGCGCTTGCATGATAGCAGATAGAGAAA 540
QY 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
DB 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAAGAGTTCCTTCTCAGGA 600
QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 601 TCTTTTGTCTATATTCTGGCTTAAGAGAGAGAGTCTTATGCGCAGGACTCATTCTTTC 660
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
DB 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGTGCTGATGTTCCAA 720
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
DB 721 TACTTAGTAAATAAGCCTTCAGAAAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAA 780

; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-1

Alignment Scores:
Pred. No.:          9.6e-217      Length:      1053
Score:              1821.00      Matches:     351
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%  Mismatches:  0
Query Match:        100.00%      Indels:      0
DB:                  18          Gaps:         0

US-10-698-228-1 (1-351) x US-10-698-228-2 (1-1053)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
DB 1 ATGGGCGACCCGGAAAGCGGAGCGCGCGCTGGATCAGATGAGATCATCTTCA 60
QY 21 AspThrAsnGluSerGluLeuLysSerAsnGluGluProLeuLeuArgLysSerArg 40
DB 61 GACACCAACGAAAGTGAATGAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
DB 121 CGGTTTGTCTATCTTCCATCCAGTACCTCCCTGATATTGGAAATGATATAACAGGCACAG 180
QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
DB 181 GCTTCCTTCTGCACAGCAGAGAGGTGCGACTTATCAAGAGATCTCCTCACTGGAACAAG 240
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
DB 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCTCACATCTTAGCCTTTTTCAGGCAGGTAT 300
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 360
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 361 CGCTGTCTTCTAGGCTTTCAAAATTCATCGAATGTTCACTCAGAGATGTACAGTTTG 420
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTATGCAATTGAA 480
QY 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
DB 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGCGCTTGCATGATAGCAGATAGAGAAA 540
QY 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
DB 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAAGAGTTCCTTCTCAGGA 600
QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 601 TCTTTTGTCTATATTCTGGCTTAAGAGAGAGAGTCTTATGCGCAGGACTCATTCTTTC 660
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
DB 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGTGCTGATGTTCCAA 720
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
DB 721 TACTTAGTAAATAAGCCTTCAGAAAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAA 780
```

```

QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaIleSerAsp 100
DB 241 CTTAAAGCAGATGAGAGTACTTCTCATCTCTACATCTTAGCCCTTTTTCAGCCAGTGAT 300
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 301 GGAATTGTAATGAAAATTGGTGGAGCGCTTTAGTCAGGAGGTGAGGTTCCAGAGGCT 360
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTG 420
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 421 CTGATAGACACTTACATCAGATCCCAAGAAAAGGGAATTTTATTTATGCAATTTGAA 480
QY 161 ThrMetProTyrValLysLysAlaAspTyrAlaLeuArgTyrIleAlaAspArgLys 180
DB 481 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGCCCTTGGATGGATAGCAGATAGAAAA 540
QY 181 SerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePheSerGly 200
DB 541 TCTACTTTTGGGAAAGAGTGGTGGCTTGTCTGTAGAGAGATTTCTTCTCAGGA 600
QY 201 SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 601 TCTTTTGTCTATATCTTGGCTTAAAGAGAGAGGCTTATGCGAGGACTCACTTTTTC 660
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
DB 661 AATGAACCTCATCAGCAGATGAAGGACTTCACTGTGACTTTTGTCTTGCCTGTGATTTCAA 720
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
DB 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGCTCAGGAGATCAATGTTGATGCTGTCTCAA 780
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
DB 781 ATTGACGAGGATTTTAAACAGAGGCTTGCCAGTTGGCCTCATTTGGAATGAATGCATT 840
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
DB 841 TTGATGAAACAGTACATTCAGTTGTAGCTGACAGATTACTTGTGAACTTGGATTTCTCA 900
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
DB 901 AAGGTTTTTTCAGCAGAAAATCCTTTGATTTTATGGAATAACATTTCTTTAGAGAGGAAA 960
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
DB 961 ACMAATTTCTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB 1021 ACAGATAACGCTCTTCACTTGGATGAGATTTT 1053

```

RESULT 3

```

US-10-698-228-4
; Sequence 4, Application US/10698228
; Publication No. US20040072253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619W00P
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770

```

```

; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 4
; LENGTH: 1081
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-4

```

```

Alignment Scores:
Pred. No.: 1e-216
Score: 1821.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
Indels: 0
Gaps: 0

```

US-10-698-228-1 (1-351) x US-10-698-228-4 (1-1081)

```

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
DB 20 ATGGCGCAGCCGGAAGGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 79
QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
DB 80 GACACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 139
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
DB 140 CGGTTTGTCTATCTTTTCCAATCCAGTACCTGATATTGGAAAAATGATAAAACAGGCACAG 199
QY 61 AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
DB 200 GCTTCCCTCTTGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCAGTGGAAACAAG 259
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
DB 260 CTTAAAGCAGATGAGAGTACTTTCATCTCTCACATCTTTAGCCTTTTTCAGCCAGTGAT 319
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 320 GGAATTGTAATGAAAATTGGTGGAGCGCTTTAGTCAGGAGGTGAGGTTCCAGAGGCT 379
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 380 CGCTGTTTCTATGGCTTTCAAAATCTCATCAGAGATGTTCACTCAGAGATGACAGTTTG 439
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 440 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA 499
QY 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTyrIleAlaAspArgLys 180
DB 500 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGCCCTTGGATGGATAGCAGATAGAAAA 559
QY 181 SerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePheSerGly 200
DB 560 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 619
QY 201 SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 620 TCTTTTGTCTATATTTCTGGCTAAGAGAGAGGCTTATGCGCAGGACTCACTTTTTC 679
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
DB 680 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 739
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
DB 740 TACTTAGTAAATAAGCCTTCAAGAAAAGGCTCAGGAGATCAATGTTGATGCTGTGCAA 799
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280

```

```
Db      800 ATTGACGAGGATTTTAAACAGAGGCTTCCAGTTGGCCCTCATTTGGAATGAATTGCATT 859
Qy      281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db      860 TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTGGATTCTCA 919
Qy      301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db      920 AAGGTTTTTCGCGCAGAAAATCCTTTTGAATTTTATGGAACAACATTTCTTTAGAGGAAAA 979
Qy      321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db      980 ACAATTTCTTTGAGAAACAGATTTTACAGATATCAGCGTTTTCAGTTTATGGCAGAAACC 1039
Qy      341 ThrAsnValPheThrLeuAspAlaAspPhe 351
Db      1040 ACAGATAAGCTCTTCACCTTGGATGCAGATTTT 1072

RESULT 4
US-10-172-118-71
; Sequence 71, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; PRIOR FILING DATE: 2002-06-14
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 71
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ENTRY NUMBER: AB036063
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-71

Alignment Scores:
Pred. No.:      11e-215      Length:      4955
Score:          1821.00      Matches:      351
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:              17          Gaps:          0

US-10-698-228-1 (1-351) x US-10-172-118-71 (1-4955)

Qy      1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20
Db      245 ATGGGGCAGCCCGGAAGCGCGGAGCGCGGCTGGATCAGATGAGATCATCTTCA 304
Qy      21 AspThrAsnGluSerGlnIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
Db      305 GACACCAACCAAGTGAATTAAGTCAATATGAAGCCACTCTCTAAGAAAGATTTCTCGC 364
Qy      41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
Db      365 CGTTTGTTCATCTTCCATCCAGTACCCTGATATTTGGAAATGTATAAACAGGCACAG 424
Qy      61 AlaSerPheThrAlaGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
Db      425 GCTTCCTCTGGACAGCAGAGAGGTCGACTTATCAAGGATCTCCCTCACTGGAAACAAG 484
```

```
Qy      81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp 100
Db      485 CTTAAACAGCATGAGAAAGTACTTCTCATCTCTCACATCTTTAGCCTTTTTCAGCCAGTGAT 544
Qy      101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db      545 GGAATTTGTAATGAAAATTTTGTGAGCGCTTTTGTACAGGAGGTCCAGGTTCCAGAGGCT 604
Qy      121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db      605 CGCTGTTTCTATGCTTCAAATTTCTCATCAGAAATGTTCACTCAGAGATGTACAGTTTG 664
Qy      141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db      665 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAAATCAATGAA 724
Qy      161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTTrpIleAlaAspArgLys 180
Db      725 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTTGCCGATGATAGCAGATGAAAA 784
Qy      181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
Db      785 TCTACTTTTGGGAAAGAGTGTGCTTCTGCTGTAGAGGAGTTTCTTCTCAGCA 844
Qy      201 SerPheAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db      845 TCTTTGCTGCTATATTCTGGCTAAAGAGAGAGGCTCTTATGCCAGGACTCACTTTTTC 904
Qy      221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db      905 AATGAATCTCATCAGCAGAGATGAAGGACTTTCACCTGTGACTTTGCTTGCTGTGATGTTCAA 964
Qy      241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
Db      965 TACTTAGTAATAAGCCTTCAGAAAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 1024
Qy      261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db      1025 ATTGAGCAGGAGTTTAAACAGAAAGCCTTGCCAGTTGGCCCTCATTTGGAATGAATTCATT 1084
Qy      281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db      1085 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTACTTTGGAACCTTGGATTTCTCA 1144
Qy      301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db      1145 AAGTTTTTTCAGGCAGAAAATCTTTTGAATTTTATGGAACAACATTTCTTTAGAGAAAA 1204
Qy      321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db      1205 ACAATTTCTTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTTTATGCGAGAAACC 1264
Qy      341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db      1265 ACAGATAACGCTCTTCACTTGGATGCAGATTTT 1297
```

RESULT 5

```
US-10-342-887-71
; Sequence 71, Application US/10342887
; Publication No. US20040059340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
```


; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 71
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-71

Alignment Scores:

Pred. No.: 1,1e-215 Length: 4955
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0

US-10-698-228-1 (1-351) x US-10-342-887-71 (1-4955)

Qy 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
Db 245 ATGGCGAGCCCGGAAAGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
Qy 21 AspThrAsnGluSerGluLeuLysSerAsnGluGluProLeuLeuArgLysSerArg 40
Db 305 GACACCAACGAAAGTAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 364
Qy 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
Db 365 CGTTTGTTCATCTTCCAAATCCAGTACCTCCCTGATATTTGGAAATGTATAACAGGCACAG 424
Qy 61 AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
Db 425 GCTTCTCTTGGACAGCAGAGAGTGCATCTTATCAAGAGATCTCCCTCACTGGAACAAG 484
Qy 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp 100
Db 485 CTTAAAGCAGATGAGAAGTACTTCACTCTCACATCTTAGCCTTTTGGAGCCAGTGAT 544
Qy 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 545 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGTTCCAGAGGCT 604
Qy 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 605 CGCTGTTTCTATGGCTTTCAATCTCATCGAATGTTCACTCAGAGATGTACAGTTTG 664
Qy 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAA 724
Qy 161 ThrMetProTyrValLysLysLysAlaAspTyrAlaLeuArgTyrIleAlaAspArgLys 180
Db 725 ACCATCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGGCATGGATGAGATAGAAAA 784
Qy 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
Db 785 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAGAGTTTCTTCTCAGGA 844
Qy 201 SerPheAlaAlaIlePheTyrLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 845 TCTTTTGTCTATATCTTGGCTAAAGACAGAGGTCTTATGCCAGGACTCACTTTTTC 904
Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 905 AATGAACATCATCAGCAGATGAAGGACTTCACTGTGACTTTGCTGCTGATGTTCCAA 964
Qy 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260

Db 965 TACTTAGTAATAAGCCCTTCAGAAAGGGTCAGGAGATCATTTGATGCTGTCAA 1024
Qy 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 1025 ATTGAGCAGGAGTTTAAACAGAGCCCTTGCAGTTGGCTCATTCGAATGAATTCATT 1084
Qy 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 1085 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGAATTTCTCA 1144
Qy 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 1145 AAGGTTTTTCAGGCGCAAAATCCTTTGATTTTATGGAACATTTCTTTAGAAAGGAAA 1204
Qy 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 1205 ACAATTTCTTTGAGAAACGAGTTTCAGATATCAGCGTTTTCAGTTATGGCAGAAACC 1264
Qy 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1265 ACAGATAACGTTCTTCACTTGGATGCAGATTTT 1297

RESULT 6

US-10-698-228-3
; Sequence 3, Application US/10698228
; Publication No. US20040072253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 3
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-3

Alignment Scores:
Pred. No.: 1,1e-215 Length: 4955
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0

US-10-698-228-1 (1-351) x US-10-698-228-3 (1-4955)

Qy 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
Db 245 ATGGCGAGCCCGGAAAGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
Qy 21 AspThrAsnGluSerGluLeuLysSerAsnGluGluProLeuLeuArgLysSerArg 40
Db 305 GACACCAACGAAAGTAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 364
Qy 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
Db 365 CGTTTGTTCATCTTTCCAATCCAGTACCTGATATTTGGAATATGTATAACAGGCACAG 424
Qy 61 AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80

```
Db 425 GCTTCCTCTTGGACACAGAGGTCGACTTATCAAGGATCTCCCTCACTGGAACAG 484
Qy 81 LeuLeuAlaAspGluYrPheIleSerHisIleLeuAlaPhePheAlaIleSerAsp 100
Db 485 CTTAAAGCAGATGAGAAGTACTTCTCTCATCTCTTACATCTTTTGGCAGCAGTGAT 544
Qy 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 545 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGAGGTTCCAGAGCT 604
Qy 121 ArgCysPheThrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 605 CGCTGTTTCTATGGCTTTCAATCTCATCGAATGTTCACTCAGAGATGTACAGTTTG 664
Qy 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 665 CTGATAGACATTTACATCAGATCCAGAACAGAAAGGAAATTTTATTAATGCAATGAA 724
Qy 161 ThrMetProTyrValLysLysLysAlaAspTyrAlaLeuArgTrpIleAlaAspArgLys 180
Db 725 ACCATCCCTATGTTTAAGAAAAAGCAGATTTGGGCTTGGATCGATAGCAGATAGAAA 784
Qy 181 SerThrPheGlyGluArgValAlaPheAlaValGluGlyValPhePheSerGly 200
Db 785 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGTGAAGAGGTTTCTTCTCAGGA 844
Qy 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 845 TCTTTTGTCTATATCTGGCTAAGAGAGAGGCTTATGCGCAGGACTCATTITTC 904
Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 905 AATGAACCTCATCAGCAGATGAAGACITTCAGTGACTTTCCTTGGCTGATGTTCCAA 964
Qy 241 TyrLeuValAsnLysProSerGluArgValArgGluIleIleValAspAlaValLys 260
Db 965 TACTTAGTAAATAAGCTTTCAGAAAGAGGCTCAGGAGATCATTTGTGATGCTGTCAA 1024
Qy 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 1025 ATTGACGAGAGTTTAAACAGAGCTTGCAGGCTTGCCTCATTTGGAATGCAATT 1084
Qy 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 1085 TTGATGAACAGTACATCAGTTGAGTTGTAGCTGACAGATTACTTGTGAACTTGGATTCTCA 1144
Qy 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 1145 AAGGTTTTTCAGGCAGAAAATCCTTTTGATTTTATGGAACAACATTTCTTTAGAGGAAA 1204
Qy 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 1205 ACAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAAGAAC 1264
Qy 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1265 ACAGATAACGTCTTCACTTGGATGAGATTTT 1297
```

RESULT 7

```
US-09-925-301-505
; Sequence 505, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
```

```
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 505
; LENGTH: 1989
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1917)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-505
```

```
Alignment Scores:
Pred. No.: 2,82e-171 Length: 1989
Score: 1460.50 Matches: 277
Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
DB: 9 Gaps: 1
```

US-10-698-228-1 (1-351) x US-09-925-301-505 (1-1989)

```
Qy 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
Db 185 CTTGAGCGGACCCGCGCTCTGGCCAGCAAGCCGAGGAGGATCTTCCAGAGGCCAC 244
Qy 23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
Db 245 GGAGCCGAAACTAAAGCAGCTGCCCCGCGGTGGAGGATGAGCGCTGTGAGAGAAA 304
Qy 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGl 58
Db 305 CCCCAGCGCTTTGTCTATCTTCCCATCGAGTACCATGATATCTGGCAGATGATAAGAA 364
Qy 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAsnLeuSerLysAspLeuProHisTr 78
Db 365 GGCAGAGGCTTCTTTTGGACCCGCGAGGAGTGAGCTCTCCAGGACATTCAGACATG 424
Qy 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAl 98
Db 425 GGAATCCTCTGAACCCGAGGAGAGATATTTATATCCATGTTCTGGCTTCTTTGACG 484
Qy 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
Db 485 AAGCGATGGCATAGTAAATGAAACTTGGTGGAGCGATTAGCCAAAGAAAGTTTCAGATT 544
Qy 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
Db 545 AGAAGCCCGCTGTTTCTATGGCTTCCAAATTTGCCATGGAAACATACATCTCTGAAATG 604
Qy 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
Db 605 TAGTCTCTTATTCACACTTACATAAAGATCCCAAGAAAGGAAATTTCTTCAATGC 664
Qy 158 alleGluThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
Db 665 CATTGAAACGATGCTTGTGTCAAGAAGAGGAGACTGGGCTTGCCTGGATTGGGA 724
Qy 178 pArgLysSerThrPheGlyGluArgValAlaPheAlaValGluGlyValPhePh 198
Db 725 CAAAGAGGCTTACCTATGTTGAACGCTGTTGTAGCCCTTTGCTCAGTGGAGGCAATTTCT 784
Qy 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
Db 785 TTCGGTCTTTTTCGTCGATATTTCTGGCTCAAGAAACGAGGACTGTATGCTTGGCTTCT 844
Qy 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
Db 845 ATTTCTTAATGAACCTTATTAGCAGAGATGAGGTTTACACTGTGATTTGCTTGCCTGAT 904
Qy 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
Db 905 GTTCAAAACCTGTGTACACAAACCATCGGAGGAGAGATTAAGAGAAATAATATTCATATGC 964
```

QY	258	aVallysIleGluGlnGlnPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs	278
		::: :::	
Db	965	TGTTGGATAGAAcAGAGGTTCTCACTGAGGCCTTGCTGTGAAGCTCATTTGGGATGA	1024
QY	278	nCySIIeLeuMetLysGlnTyrlleGluPheValAlaAspArgLeuLeuValGluLeuG	298
Db	1025	TTGCACCTCTAATGAAGCAATACATTGAGTTTGTGGCAGACAGACTTATTCTGGAACCTGGG	1084
QY	298	yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG	318
Db	1085	TTTTAGCAAGTTTTCAGAGTAGAAGACCATTGACTTTATGGAGAATATTTCACTGGA	1144
QY	318	uGlyLysThrAsnPhePheGluLysArgValSerGluTyrlGlnArgPheAlaValMetAl	338
Db	1145	AGGAAGACTCACTTCTTTGAGAGAGAGTAGCGGAGTATCAGAGGATGGGAGTGATGTC	1204
QY	338	aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe	351
		: ::: :::	
Db	1205	AACTCCAAcACAGAAATCTTTTACCTTGGATGCTGACTTC	1244

RESULT 8

```

US-10-084-817-342
; Sequence 342, Application US/10084817
; Publication No. US20030119009A1
; GENERAL INFORMATION:
; APPLICANT: Susan Stuart
; APPLICANT: Jed G. Nuchtern
; APPLICANT: Sharon E. Plon
; APPLICANT: Jason M. Shohet
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
; FILE REFERENCE: PA-0046 US
; CURRENT APPLICATION NUMBER: US/10/084,817
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/270,784
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 365
; SOFTWARE: PERL Program
; SEQ ID NO 342
; LENGTH: 2216
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030119009A1 202239.1
US-10-084-817-342

```

[illegible]

US-10-698-228-1 (1-351) x US-10-084-817-342 (1-2216)

QY	4	ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs	23
DB	189	CCTGAGCGGACCGCGCTCTCTGGCCAGCAGACCGCAGAGGATCTTCCAGGAGCCAC	248
QY	23	nGluSerGluIleIleIysSer-----AsnGluGluProLeuLeuArgIysSe	38
DB	249	GGAGCGGAAACTAAAGCAGCGTGGCGCGGTGGAGGATGAGCGCTCTCAGAGAGAAA	308
QY	38	rSerArgArgPheValIlePheProIleGlnTrpProAspIleTrpLysMetTrpLysGl	58
DB	309	CCCCCGCGCTTGTTCATCTTCCCATCGAGTACCATGATATCTCGCAGATGTATAGAA	368
QY	58	nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTr	78
DB	369	GGCAGAGGCTTCCTTTTGGACCGCGGAGAGGTGACCTCTCCAAGGACATTCAGCATGT	428

RESULT 9

```

US-09-954-456-724
; Sequence 724, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25

```

; PRIOR APPLICATION NUMBER: US/60/235,134
 ; PRIOR FILING DATE: 2000-09-25
 ; PRIOR APPLICATION NUMBER: US/60/235,637
 ; PRIOR FILING DATE: 2000-09-26
 ; PRIOR APPLICATION NUMBER: US/60/235,638
 ; PRIOR FILING DATE: 2000-09-26
 ; PRIOR APPLICATION NUMBER: US/60/235,711
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: US/60/235,720
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: US/60/235,840
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: US/60/235,863
 ; PRIOR FILING DATE: 2000-09-27
 ; NUMBER OF SEQ ID NOS: 2276
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 724
 ; LENGTH: 2500
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-954-456-724

Alignment Scores:
 Pred. No.: 4,04e-171 Length: 2500
 Score: 1460.50 Matches: 277
 Percent Similarity: 88.42% Conservative: 36
 Best Local Similarity: 78.25% Mismatches: 35
 Query Match: 80.20% Indels: 6
 DB: 9 Gaps: 1

US-10-698-228-1 (1-351) x US-09-954-456-724 (1-2500)

QY 4 ProGluArgProGluAlaAlaGluLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
 DB 302 CCTGACGGGACCCGGCTGCTGCCACAGCGGAGGAGGATCTTCAGAGGCCAC 361
 QY 23 nGluSerGluLeuLysSer-----AsnGluGluProLeuLeuArgLysSe 38
 DB 362 GGAGCCGAAACTAAAGACAGCTGCCCGCGGTGGAGGATGAGCGCTGCTGAGAGAAA 421
 QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysG 58
 DB 422 CCCCCCGCGCTTGTATCTTCCTCCATCGAGTACCATGATATCTGCAGATGTATAGAA 481
 QY 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAlaLeuSerLysAspLeuProHisTr 78
 DB 482 GGCAGAGGCTTCTTTTGGACCCCGAGGAGGTGACCTCTCCAGACATTCAGCACTG 541
 QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAl 98
 DB 542 GGAATCCCTGAAACCCGAGGAGAGATATTTATATCCCATGTTCTGGCTTTCTTCGAGC 601
 QY 98 aSerAspGlyIleValAsnGlnAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
 DB 602 AAGCGATGCATAGTAATAAGTAACTTGTGGAGCGATTTAGCAAGAAAGTTTCAGATTAC 661
 QY 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
 DB 662 AGAAGCCCGCTGTTCTATGGCTTCAAAATGGCATGCAAAACATCATCTCTGAATGTA 721
 QY 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
 DB 722 TAGCTCTTATTGACACTTACATAAAGATCCCAAGAAAGGAGGAAATTTCTCTCAATGC 781
 QY 158 alLeGluThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
 DB 782 CATTAAGACGATCCCTTGTGTCAAGAAGAGGAGGAGCTGGCGCTTGGCTGAGTGGGA 841
 QY 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePh 198
 DB 842 CAAAGAGGCTACCTATGTGGTGAACGTGTGTAGCTTTGTGCTGAGTGGAGGCAATTTCTT 901
 QY 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218

DB 902 TTCGGGTCTTTTGGCGCATATTCTGGCTCAAGAAACGAGGACTGATGCTGGCTCAC 961
 QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
 DB 962 ATTTCTTAATGAACCTATTAGCAGAGATGAGGTTTACACTGTGATTTGCTTGCCTGAT 1021
 QY 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluLeuIleValAspAl 258
 DB 1022 GTTCAAAACACCTGGTACACAAACCATCGAGGAGAGTAGAGAAATAATTATCAATGC 1081
 QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
 DB 1082 TGTTGGATAGAACAGAGTTCTCTACTGAGGCTTGGCTGTGAAGCTCATTTGGATGAA 1141
 QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuG 298
 DB 1142 TTGCACCTTAATGAGCAATACATTGAGTTTGTGGCAGACAGACTTATGCTGGAACTGG 1201
 QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG 318
 DB 1202 TTTTAGCAGGTTTTCAGAGTAGAGAACCCATTGACTTTATGAGAAATAATTCACCTGGA 1261
 QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
 DB 1262 AGGAAAGACTAATCTTTCAGAGAGAGAGTAGGCGCGATTCAGAGGATGGAGTGTATGTC 1321
 QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
 DB 1322 AAGTCCACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1361

RESULT 10

; US-09-954-456-1169
 ; Sequence 1169, Application US/09954456
 ; Patent No. US20020115057A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Young, Paul
 ; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
 ; TITLE OF INVENTION: Sets
 ; FILE REFERENCE: 689290-76
 ; CURRENT APPLICATION NUMBER: US/09/954,456
 ; CURRENT FILING DATE: 2001-09-18
 ; PRIOR APPLICATION NUMBER: US/60/233,617
 ; PRIOR FILING DATE: 2000-09-18
 ; PRIOR APPLICATION NUMBER: US/60/234,052
 ; PRIOR FILING DATE: 2000-09-20
 ; PRIOR APPLICATION NUMBER: US/60/234,923
 ; PRIOR FILING DATE: 2000-09-25
 ; PRIOR APPLICATION NUMBER: US/60/235,134
 ; PRIOR FILING DATE: 2000-09-25
 ; PRIOR APPLICATION NUMBER: US/60/235,637
 ; PRIOR FILING DATE: 2000-09-26
 ; PRIOR APPLICATION NUMBER: US/60/235,638
 ; PRIOR FILING DATE: 2000-09-26
 ; PRIOR APPLICATION NUMBER: US/60/235,711
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: US/60/235,720
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: US/60/235,840
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: US/60/235,863
 ; PRIOR FILING DATE: 2000-09-27
 ; NUMBER OF SEQ ID NOS: 2276
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 1169
 ; LENGTH: 2500
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-954-456-1169

Alignment Scores: 4.04e-171 Length: 2500
 Pred. No.: 1460.50 Matches: 277
 Score:

Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
DB: 9 Gaps: 1

US-10-698-228-1 (1-351) x US-09-954-456-1169 (1-2500)

```
QY 4 ProGluArgProGluAlaAlaGlyLeuAspGluArgSerSerSerAspThrAs 23
DB 302 CCTGAGCGGACCCGGCTCTGCGCAGCAAGACCGCGAGGAGTCTTCCAGGAGCCAC 361
QY 23 nGluSerGluLeuSer-----AsnGluGluProLeuLeuArgLysSe 38
DB 362 GGAGCCGAAACTAAAGCAGCTCCCGCGCGTGGAGGATGAGCGCTGCTGAGAGAAA 421
QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysG1 58
DB 422 CCCCCCGCGCTTGTCTATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481
QY 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
DB 482 GGCAGAGGCTCTCTTTGGACCGCGGAGGAGTTGACCTCTCCAAGGACATTCAGCACTG 541
QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAl 98
DB 542 GGAATCCCTGAACCCGCGAGGAGATATTTATATCCATGTTCTTGGCTTCTTTCGAGC 601
QY 98 aSerAspGlyIleValAsnGluLeuValGluArgPheSerGlnGluValGlnValPr 118
DB 602 AGCGGATGGCATAGTAATAAGAAACTTGGTGGAGCGATTTAGCCAAAGATTCAGATTAC 661
QY 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
DB 662 AGAAGCGCGCTGTTCTATGGCTTCCAAATTCGATGCAAAACATACATTTCTGAATGTA 721
QY 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
DB 722 TAGCTCTCTTATTGACACTTACATAAAGATCCCAAGAAAGGAATTTCTTCTTCAATGC 781
QY 158 alIeGluThrMetProTyrValLysLysAlaAspTrpAlaIleuArgTrpIleAlaAs 178
DB 782 CATTTGAACGATGCCCTTGTGTCAAGAAAGGAGGAGCTGGCGCTTGGCTGGATGGGA 841
QY 178 pArgLysSerThrPheGluGluArgValValAlaPheAlaAlaValGluGluValPhePh 198
DB 842 CAAAGAGGCTACCTATGGTGAACGTTGTGAGCTTGTGCGAGTGAAGCATTTCTT 901
QY 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
DB 902 TTCCGGTCTTTTGGCTCGATATTTCTGGCTCAAGAAACGAGGACTGATGCTGGCCTCAC 961
QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
DB 962 ATTTCTTAATGAATTAATAGCAGATGAGGGTTTACACTGTGATTTTCTGCTGCTGAT 1021
QY 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
DB 1022 GTTCAAAACCTGGTACACAAACCATCGGAGGAGAGATGAAGAAATAATATCAATGC 1081
QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValIcylLeuIleGlyMetAs 278
DB 1082 TGTTCGGATAGAACAGAGTTCTCTACTGAGGCTTGGCTGTGAAGCTCATTTGGATGAA 1141
QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuG1 298
DB 1142 TTGCACCTCTAAATGAAGCAATACATTGAGTTTGTGGCAGACAGACTTATGCTGGAACCTGG 1201
QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG1 318
DB 1202 TTTTAGCAGGTTTTCAGAGTAGAACCCATTTGACTTTATGGAGAAATATTTCACTGGA 1261
QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
```

```
DB 1262 AGGAAGACTAACTTCTTTGAGAAGAGTAGGCGAGTATCAGAGGATGGAGTGATGTC 1321
QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB 1322 AAGTCCACACAGAGAATCTTTTACCTTGGATGCTGACTTC 1361
RESULT 11
US-09-954-456-1827
; Sequence 1827, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Can
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 1827
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-1827
Alignment Scores:
Pred. No.: 4,04e-171 Length: 2500
Score: 1460.50 Matches: 277
Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
DB: 9 Gaps: 1
US-10-698-228-1 (1-351) x US-09-954-456-1827 (1-2500)
QY 4 ProGluArgProGluAlaAlaGlyLeuAspGluArgSerSerSerAspThrAs 23
DB 302 CCTGAGCGGACCCGGCTCTGCGCAGCAAGACCGCGAGGAGTCTTCCAGGAGCCAC 361
QY 23 nGluSerGluLeuSer-----AsnGluGluProLeuLeuArgLysSe 38
DB 362 GGAGCCGAAACTAAAGCAGCTCCCGCGCGTGGAGGATGAGCGCTGCTGAGAGAAA 421
QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysG1 58
DB 422 CCCCCCGCGCTTGTCTATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481
QY 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
DB 482 GGCAGAGGCTCTCTTTGGACCGCGGAGGAGTTGACCTCTCCAAGGACATTCAGCACTG 541
QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAl 98
```

Db 542 GGATCCCTGAAACCCGAGAGAGATATTTATATCCATGTTCTGGCTTTCTTTCAGC 601
Qy 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
Db 602 AAGCGATGCGATAGTAATGAAATCTGGTGGAGCGATTAGCCAGAAAGTTCCAGATTAC 661
Qy 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
Db 662 AGAAGCCCGCTGTTCTATGGCTTCCAAATTCATGCGATGGAACATACATCTCAATGTA 721
Qy 138 rSerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 158
Db 722 TAGTCTCTTATGACACTTACATAAAGATCCCAAGAAAGGAATTTCTCTCAATGC 781
Qy 158 alleGluThrMetProTyrValLysLysLysLysLysLysLysLysLysLysLysLys 178
Db 782 CATTTGAAACGATGCTTGTCTCAAGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 841
Qy 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePh 198
Db 842 CAAAGAGGCTACCTATGCTGAACTGTTGAGGCTTTGCTGCACTGGAAGGCAATTTCTT 901
Qy 198 eSerGlySerPheAlaAlaTlePheTleLeuLeuLysLysLysLysLysLysLysLys 218
Db 902 TTCGGTCTTTTGGCTGATATTTCTGCTCAAGAAACGAGGAGGAGGAGGAGGAGGAG 961
Qy 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
Db 962 ATTTCTATGACTTATTAGCAGAGATGAGGGTTTACACTGTGATTTGCTTGCCTGAT 1021
Qy 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
Db 1022 GTTCAAAACACCTGTTACACAAACCATCGGAGGAGAGAGTAAAGAAATATTTCAATGC 1081
Qy 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
Db 1082 TGTTCGGATAGAACAGAGGTTCTCTACTGAGGCTTGGCTGTGAAGCTCAATGGAGTAA 1141
Qy 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuG 298
Db 1142 TTGCACCTTAATGAGCAATACATGATTGTTGTCAGACAGACTTATGCTGGAATGGG 1201
Qy 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG 318
Db 1202 TTTTAGCAAGGTTTTCAGAGTAGAGAACCCATTTGACTTTATGGAGAAATATTTCACTGGA 1261
Qy 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
Db 1262 AGGAAAGACTAACTTCTTTGAGAGAGAGTAGCGGAGTATCAGAGGATGGGAGTGTGTC 1321
Qy 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1322 AGTCCAAACAGAAATCTTTTACCTTGGATGCTGACTTC 1361

RESULT 12

US-10-641-643-1370

; Sequence 1370, Application US/10641643

; Publication No. US20040077003A1

; GENERAL INFORMATION:

; APPLICANT: Cocks, Benjamin G.

; Suan G. Stuart

; Jeffrey J. Seilhamer

; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL

; NUMBER OF SEQUENCES: 1508

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

; STREET: 3174 PORTER DRIVE

; CITY: PALO ALTO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; 158 alleGluThrMetProTyrValLysLysLysLysLysLysLysLysLysLysLysLys 178

; 138 rSerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 158

; 662 AGAAGCCCGCTGTTCTATGGCTTCCAAATTTGCGATGGAACATACATTTCTGAAATGTA 721

; 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138

; 602 AAGCGATGCGATAGTAATGAAACTTGGTGGAGCGATTAGCCAAAGAAAGTTCCAGATTAC 661

; 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118

; 542 GGAATCCCTGAAACCCGAGGAGAGATATTTATATCCCATGTTCTGGCTTTCTTTCAGC 601

; 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAl 98

; 482 GGCAGAGGCTTCTTTTGGACCCGAGGAGGTTGACCTCTCCAAGGACATTCAGCAGCTG 541

; 58 nAlaGluAlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78

; 422 CCCCGCGCTTGTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481

; 38 rSerArgArgPheValIlePhePheProfileGlnTyrProAspIleTrpLysMetTyLysG 58

; 23 nGluSerGluLysSer-----AsnGluGluProLeuLeuArgLysSe 38

; 302 CCTGAGCGGACCCGCGCTCTGCGCCAGACACCGCGAGGAGGATCTTCCAGAGGCCAC 361

; 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23

; 422 CCCCGCGCTTGTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481

; 362 GGAGCCGAAACTAAAGCAGCTGCCCCCGCGGTGGAGGATGAGCGCTGCTGAGAGAAA 421

; 58 nAlaGluAlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78

; 422 CCCCGCGCTTGTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481

; 38 rSerArgArgPheValIlePhePheProfileGlnTyrProAspIleTrpLysMetTyLysG 58

; 23 nGluSerGluLysSer-----AsnGluGluProLeuLeuArgLysSe 38

; 302 CCTGAGCGGACCCGCGCTCTGCGCCAGACACCGCGAGGAGGATCTTCCAGAGGCCAC 361

; 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23

; 422 CCCCGCGCTTGTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481

; 362 GGAGCCGAAACTAAAGCAGCTGCCCCCGCGGTGGAGGATGAGCGCTGCTGAGAGAAA 421

; 58 nAlaGluAlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78

; 422 CCCCGCGCTTGTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481

; 38 rSerArgArgPheValIlePhePheProfileGlnTyrProAspIleTrpLysMetTyLysG 58

; 23 nGluSerGluLysSer-----AsnGluGluProLeuLeuArgLysSe 38

; 302 CCTGAGCGGACCCGCGCTCTGCGCCAGACACCGCGAGGAGGATCTTCCAGAGGCCAC 361

; 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23

; 422 CCCCGCGCTTGTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481

; 362 GGAGCCGAAACTAAAGCAGCTGCCCCCGCGGTGGAGGATGAGCGCTGCTGAGAGAAA 421

; 58 nAlaGluAlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78

; 422 CCCCGCGCTTGTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481

; 38 rSerArgArgPheValIlePhePheProfileGlnTyrProAspIleTrpLysMetTyLysG 58

; 23 nGluSerGluLysSer-----AsnGluGluProLeuLeuArgLysSe 38

; 302 CCTGAGCGGACCCGCGCTCTGCGCCAGACACCGCGAGGAGGATCTTCCAGAGGCCAC 361

; 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23

; 422 CCCCGCGCTTGTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481

; 362 GGAGCCGAAACTAAAGCAGCTGCCCCCGCGGTGGAGGATGAGCGCTGCTGAGAGAAA 421

; 58 nAlaGluAlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78

; 422 CCCCGCGCTTGTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481

; 38 rSerArgArgPheValIlePhePheProfileGlnTyrProAspIleTrpLysMetTyLysG 58

; 23 nGluSerGluLysSer-----AsnGluGluProLeuLeuArgLysSe 38

; 302 CCTGAGCGGACCCGCGCTCTGCGCCAGACACCGCGAGGAGGATCTTCCAGAGGCCAC 361

; 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23

; 422 CCCCGCGCTTGTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481

; 362 GGAGCCGAAACTAAAGCAGCTGCCCCCGCGGTGGAGGATGAGCGCTGCTGAGAGAAA 421

; 58 nAlaGluAlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78

; 422 CCCCGCGCTTGTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481

; 38 rSerArgArgPheValIlePhePheProfileGlnTyrProAspIleTrpLysMetTyLysG 58

; 23 nGluSerGluLysSer-----AsnGluGluProLeuLeuArgLysSe 38

; 302 CCTGAGCGGACCCGCGCTCTGCGCCAGACACCGCGAGGAGGATCTTCCAGAGGCCAC 361

; 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23

; 422 CCCCGCGCTTGTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481

; 362 GGAGCCGAAACTAAAGCAGCTGCCCCCGCGGTGGAGGATGAGCGCTGCTGAGAGAAA 421

; 58 nAlaGluAlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78

; 422 CCCCGCGCTTGTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481

; 38 rSerArgArgPheValIlePhePheProfileGlnTyrProAspIleTrpLysMetTyLysG 58

; 23 nGluSerGluLysSer-----AsnGluGluProLeuLeuArgLysSe 38

; 302 CCTGAGCGGACCCGCGCTCTGCGCCAGACACCGCGAGGAGGATCTTCCAGAGGCCAC 361

; 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23

; 422 CCCCGCGCTTGTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481

; 362 GGAGCCGAAACTAAAGCAGCTGCCCCCGCGGTGGAGGATGAGCGCTGCTGAGAGAAA 421

; 58 nAlaGluAlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78

; 422 CCCCGCGCTTGTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481

; 38 rSerArgArgPheValIlePhePheProfileGlnTyrProAspIleTrpLysMetTyLysG 58

; 23 nGluSerGluLysSer-----AsnGluGluProLeuLeuArgLysSe 38

; 302 CCTGAGCGGACCCGCGCTCTGCGCCAGACACCGCGAGGAGGATCTTCCAGAGGCCAC 361

; 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23

; 422 CCCCGCGCTTGTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481

; 362 GGAGCCGAAACTAAAGCAGCTGCCCCCGCGGTGGAGGATGAGCGCTGCTGAGAGAAA 421

; 58 nAlaGluAlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78

; 422 CCCCGCGCTTGTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481

; 38 rSerArgArgPheValIlePhePheProfileGlnTyrProAspIleTrpLysMetTyLysG 58

; 23 nGluSerGluLysSer-----AsnGluGluProLeuLeuArgLysSe 38

; 302 CCTGAGCGGACCCGCGCTCTGCGCCAGACACCGCGAGGAGGATCTTCCAGAGGCCAC 361

; 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23

; 422 CCCCGCGCTTGTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481

; 362 GGAGCCGAAACTAAAGCAGCTGCCCCCGCGGTGGAGGATGAGCGCTGCTGAGAGAAA 421

; 58 nAlaGluAlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78

; 422 CCCCGCGCTTGTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481

; 38 rSerArgArgPheValIlePhePheProfileGlnTyrProAspIleTrpLysMetTyLysG 58

; 23 nGluSerGluLysSer-----AsnGluGluProLeuLeuArgLysSe 38

; 302 CCTGAGCGGACCCGCGCTCTGCGCCAGACACCGCGAGGAGGATCTTCCAGAGGCCAC 361

; 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23

; 422 CCCCGCGCTTGTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481

; 362 GGAGCCGAAACTAAAGCAGCTGCCCCCGCGGTGGAGGATGAGCGCTGCTGAGAGAAA 421

; 58 nAlaGluAlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78

; 422 CCCCGCGCTTGTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481

; 38 rSerArgArgPheValIlePhePheProfileGlnTyrProAspIleTrpLysMetTyLysG 58

; 23 nGluSerGluLysSer-----AsnGluGluProLeuLeuArgLysSe 38

; 302 CCTGAGCGGACCCGCGCTCTGCGCCAGACACCGCGAGGAGGATCTTCCAGAGGCCAC 361

; 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23

; 422 CCCCGCGCTTGTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481

; 362 GGAGCCGAAACTAAAGCAGCTGCCCCCGCGGTGGAGGATGAGCGCTGCTGAGAGAAA 421

; 58 nAlaGluAlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78

; 422 CCCCGCGCTTGTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481

; 38 rSerArgArgPheValIlePhePheProfileGlnTyrProAspIleTrpLysMetTyLysG 58

; 23 nGluSerGluLysSer-----AsnGluGluProLeuLeuArgLysSe 38

; 302 CCTGAGCGGACCCGCGCTCTGCGCCAGACACCGCGAGGAGGATCTTCCAGAGGCCAC 361

; 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23

; 422 CCCCGCGCTTGTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481

; 362 GGAGCCGAAACTAAAGCAGCTGCCCCCGCGGTGGAGGATGAGCGCTGCTGAGAGAAA 421

; 58 nAlaGluAlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78

; 422 CCCCGCGCTTGTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481

; 38 rSerArgArgPheValIlePhePheProfileGlnTyrProAspIleTrpLysMetTyLysG 58

; 23 nGluSerGluLysSer-----AsnGluGluProLeuLeuArgLysSe 38

; 302 CCTGAGCGGACCCGCGCTCTGCGCCAGACACCGCGAGGAGGATCTTCCAGAGGCCAC 361

; 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23

; 422 CCCCGCGCTTGTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481

; 362 GGAGCCGAAACTAAAGCAGCTGCCCCCGCGGTGGAGGATGAGCGCTGCTGAGAGAAA 421

; 58 nAlaGluAlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78

; 422 CCCCGCGCTTGTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481

; 38 rSerArgArgPheValIlePhePheProfileGlnTyrProAspIleTrpLysMetTyLysG 58

; 23 nGluSerGluLysSer-----AsnGluGluProLeuLeuArgLysSe 38

; 302 CCTGAGCGGACCCGCGCTCTGCGCCAGACACCGCGAGGAGGATCTTCCAGAGGCCAC 361

; 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23

; 422 CCCCGCGCTTGTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481

; 362 GGAGCCGAAACTAAAGCAGCTGCCCCCGCGGTGGAGGATGAGCGCTGCTGAGAGAAA 421

; 58 nAlaGluAlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78

; 422 CCCCGCGCTTGTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481

; 38 rSerArgArgPheValIlePhePheProfileGlnTyrProAspIleTrpLysMetTyLysG 58

; 23 nGluSerGluLysSer-----AsnGluGluProLeuLeuArgLysSe 38

; 302 CCTGAGCGGACCCGCGCTCTGCGCCAGACACCGCGAGGAGGATCTTCCAGAGGCCAC 361

; 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23

; 422 CCCCGCGCTTGTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481

; 362 GGAGCCGAAACTAAAGCAGCTGCCCCCGCGGTGGAGGATGAGCGCTGCTGAGAGAAA 421

; 58 nAlaGluAlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78

; 422 CCCCGCGCTTGTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481

; 38 rSerArgArgPheValIlePhePheProfileGlnTyrProAspIleTrpLysMetTyLysG 58

; 23 nGluSerGluLysSer-----AsnGluGluProLeuLeuArgLysSe 38

; 302 CCTGAGCGGACCCGCGCTCTGCGCCAGACACCGCGAGGAGGATCTTCCAGAGGCCAC 361

; 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23

; 422 CCCCGCGCTTGTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481

; 362 GGAGCCGAAACTAAAGCAGCTGCCCCCGCGGTGGAGGATGAGCGCTGCTGAGAGAAA 421

; 58 nAlaGluAlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78

; 422 CCCCGCGCTTGTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481

; 38 rSerArgArgPheValIlePhePheProfileGlnTyrProAspIleTrpLysMetTyLysG 58

; 23 nGluSerGluLysSer-----AsnGluGluProLeuLeuArgLysSe 38

; 302 CCTGAGCGGACCCGCGCTCTGCGCCAGACACCGCGAGGAGGATCTTCCAGAGGCCAC 361

; 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23

; 422 CCCCGCGCTTGTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481

; 362 GGAGCCGAAACTAAAGCAGCTGCCCCCGCGGTGGAGGATGAGCGCTGCTGAGAGAAA 421

; 58 nAlaGluAlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78

; 422 CCCCGCGCTTGTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481

; 38 rSerArgArgPheValIlePhePheProfileGlnTyrProAspIleTrpLysMetTyLysG 58

; 23 nGluSerGluLysSer-----AsnGluGluProLeuLeuArgLysSe 38

; 302 CCTGAGCGGACCCGCGCTCTGCGCCAGACACCGCGAGGAGGATCTTCCAGAGGCCAC 361

; 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23

; 422 CCCCGCGCTTGTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481

; 362 GGAGCCGAAACTAAAGCAGCTGCCCCCGCGGTGGAGGATGAGCGCTGCTGAGAGAAA 421

```

Db      782 CATTGAACGATGCTTGTCTCAAGAAAGCGAGACTGGCCCTTGGCTGTGATGGGA 841
Qy      178 pArgLysSerThrPheGlyGluArgValAlaPheAlaValGluGlyValPhePh 198
Db      842 CAAAGAGCTACCTATGTTGTAACGTTGTGTAGCCTTGTGCTGCAAGGCAATTTCTT 901
Qy      198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
Db      902 TTCGGTTCCTTTTTCGGTCGATATTTCTGGCTCAAGAAACGAGGACTGATGCTGCGCTCAC 961
Qy      218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
Db      962 ATTTTCTAATGAACCTTATTAGCAGAGATGAGGGTTTACACTGTGATTTGCTGCCCTGAT 1021
Qy      238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleValAspAl 258
Db      1022 GTTCAACACCTGGTACACAAACCATCGAGGAGAGATGAGAGAAATATTATCAATGC 1081
Qy      258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
Db      1082 TGTTCGGATAGAACAGAGTTCTCTCACTGAGGCCCTTGCCTGTGAAGCTCATTTGGATGAA 1141
Qy      278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuGlu 298
Db      1142 TTGCACCTCTAATGAACCAATACATTTGAGTTTGTGGCAGACAGACTTATGCTGGAACCTGG 1201
Qy      298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGl 318
Db      1202 TTTTAGCAAGGTTTTCAGATAGAACCCATTTGACTTTATGGAGAAATATTTTCACTGGA 1261
Qy      318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
Db      1262 AGGAAGACTAACTTCTTTGAGAAGAGAGTAGGCGAGTATCAGAGGATGGAGTGTGTC 1321
Qy      338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db      1322 AAGTCCAAACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1361

```

RESULT 13

```

US-10-733-878-458
; Sequence 458, Application US/10733878
; Publication No. US20040224408A1
; GENERAL INFORMATION:
; APPLICANT: Jean-Philippe Girard
; APPLICANT: Francois Amalric
; APPLICANT: Myriam Rousseigne
; APPLICANT: Thomas Clouaire
; TITLE OF INVENTION: THAP PROTEINS AS NUCLEAR RECEPTORS FOR
; TITLE OF INVENTION: CHEMOKINES AND ROLES IN TRANSCRIPTIONAL REGULATION, CELL
; TITLE OF INVENTION: PROLIFERATION AND CELL DIFFERENTIATION
; FILE REFERENCE: BIOBANK.012A
; CURRENT APPLICATION NUMBER: US/10733.878
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 60/432699
; PRIOR FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/485027
; PRIOR FILING DATE: 2003-07-03
; NUMBER OF SEQ ID NOS: 535
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 458
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-733-878-458

```

Alignment Scores:

```

Pred. No.: 4.04e-171 Length: 2500
Score: 1460.50 Matches: 277
Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
Gaps: 20

```

```

US-10-698-228-1 (1-351) x US-10-733-878-458 (1-2500)
Qy      4 ProGluArgProGluAlaAlaGlyLeuAspGlu-ArgSerSerSerAspThrAs 23
Db      302 CCTGAGCGGACCCGCGCTCTGGCCAGCAAGACCGCGAGGAGGATCTTCCAGGAGCCAC 361
Qy      23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
Db      362 GGAGCCGAAACTAAAGCAGCATGCCCCCGCGGTGGAGGATGAGCGCTCTCGAGAGAAA 421
Qy      38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGl 58
Db      422 CCCCCCGCGCTTGTGCATCTTCCCATCGAGTACCATGATATCTGCAGCATGTATAGAA 481
Qy      58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
Db      482 GGCAGAGCTTCTTTTGGACCGCGAGGAGGTTCACCTCTCCAGGACATTCAGCACTG 541
Qy      78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAl 98
Db      542 GGAATCCCTGAAACCCGAGGAGAGATATTTATATATCCCATGTTCTGGCTTTCTTTCGAC 601
Qy      98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
Db      602 AAGCGATGSCATAGTAAATGAAACTTGTGTGGAGCGATTTAGCCCAAGAGATTCAGATTAC 661
Qy      118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
Db      662 AGAAGCCCGCTTGTCTATGCTTCCAAATTTGCCATGGAAAACATACATCTTGAATGT 721
Qy      138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
Db      722 TAGCTTCTTATTGACACTTACATAAAGATCCCAAGAAAGGAAATTTCTTCAATGC 781
Qy      158 aIleGluThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
Db      782 CATTGAAACGATGCTTGTGTCAAGAAAGGAGCAGACTGGGCTTTCGCTGGATTGGGGA 841
Qy      178 pArgLysSerThrPheGlyGluArgValAlaPheAlaValGluGlyValPhePh 198
Db      842 CAAAGAGCTACCTATGTTGTAACGTTGTGTAGCCTTGTGCTGAGTGAAGGCAATTTCTT 901
Qy      198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
Db      902 TTCGGTTCCTTTTTCGGTCGATATTTCTGGCTCAAGAAACGAGGACTGATGCTGCGCTCAC 961
Qy      218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
Db      962 ATTTTCTAATGAACCTTATTAGCAGAGATGAGGGTTTACACTGTGATTTTGTGCTGCTGAT 1021
Qy      238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleValAspAl 258
Db      1022 GTTCAACACCTGGTACACAAACCATCGAGGAGAGATGAGAGAAATATTATCAATGC 1081
Qy      258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
Db      1082 TGTTCGGATAGAACAGAGTTCCTCACTGAGGCCCTTGTGCTGAGCTCATTTGGGATGAA 1141
Qy      278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuGluLeuGl 298
Db      1142 TTGCACCTCTAATGAACCAATACATTTGAGTTTGTGGCAGACAGACTTATGCTGGAATGAA 1141
Qy      298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGl 318
Db      1202 TTTTAGCAAGGTTTTCAGATAGAACCCATTTGACTTTATGGAGAAATATTTTCACTGGA 1261
Qy      318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
Db      1262 AGGAAGACTAACTTCTTTGAGAAGAGAGTAGGCGAGTATCAGAGGATGGAGTGTGTC 1321
Qy      338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351

```

```
Db 1322 AAGTCCACAGAGAATCTTTTACCTTGGATGCTGACTTC 1361
RESULT 14
US-10-843-641A-3751
; Sequence 3751, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals, Inc.
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; FILE OF INVENTION: Signature Gene Sets
; FILE REFERENCE: 689290-189
; CURRENT APPLICATION NUMBER: US/10/843,641A
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/873,367
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US/09/954,531
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,832
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/964,824
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/09/967,768
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/09/968,007
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,708
; PRIOR FILING DATE: 2001-10-03
; Remaining prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 8447
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3751
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-843-641A-3751
Alignment Scores:
Pred. No.: 4,04e-171 Length: 2500
Score: 1460.50 Matches: 277
Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
DB: 21 Gaps: 1
US-10-698-228-1 (1-351) x US-10-843-641A-3751 (1-2500)
QY 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
Db 302 CCTGAGCGGAGCCCGCGTCTGCGCCAGAACAGACCGCGGAGGAGGATCTTCCAGGAGCCAC 361
QY 23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
Db 362 GGAGCGGAAACTAAGACGAGCTGCCCGCGGTGGAGGATGACCGCTGCTGAGAGAA 421
QY 38 rSerArgArgPheValIlePheProIlePheProAspIleTrpLysMetTyLysG1 58
Db 422 CCGCGCGCGCTTGTCTATCTTCCCATCGAGTACCATGATATCTGGCAGATGATAGAA 481
QY 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
Db 482 GGCAGAGCGCTTCTTTTGGACCGCGGAGGAGGTGACCTCTCCAGGACATTCAGCACTG 541
QY 78 pAsnLysLeuLysAlaAspGluLysTyPheIleSerHisIleLeuAlaPheAlaAl 98
Db 542 GGAATCCCTGAAACCCGAGGAGAGATATTTATATCCATGTTCTGCTTCTTTCGAGC 601
QY 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
Db 602 AAGCGATGGCATAGTAATAAGAAACTTGGTGAGCGGATTTAGCCAAAGATTTCAGATTAC 661
QY 118 oGluAlaArgCysPheTyrglyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
Db 662 AGAAGCCCGCTGTTTCTATGGCTTCCAATTTGCCATGGAAACATACATTTCTGAATGTA 721
QY 138 rSerLeuLeuIleAspThrTyrlleArgAspProLysLysArgGluPheLeuPheAsnAl 158
Db 722 TAGTCTTCTTATTACACTTACATAAAAGATCCCAAGAAAGGAATTTCTCTTCAATGC 781
QY 158 aIleGluThrMetProTyrvallLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
Db 782 CATTGAAACCATGCTTGTGTCAAGAAAGAGCGAGACTGGGCGCTTGGCCTGGAATTGGGA 841
QY 178 pArgLysSerThrPheGlyGluArgValAlaPheAlaAlaValGluGluValPhePh 198
Db 842 CAAAGAGGCTACCTATGCTGACCGTGTGTAGCTTGTGCTGAGTGGAGGCAATTTCTT 901
QY 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
Db 902 TTCGGTCTTCTTTCGTCGATATTTCTGGCTCAAGAAACGAGGACTGATGCTTGGCTCAC 961
QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
Db 962 ATTTTCTTAATGAACCTTATTAGCAGATGAGGGTTTACACTGTGATTTTGTCTGCTGAT 1021
QY 238 tPheGlnTyrlleValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
Db 1022 GTTCAACACCTGGTACACAAACCATCGAGGAGAGTAGAGAGAAATATTAATATCAATGC 1081
QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
Db 1082 TGTTCGATAGAACAGGAGTTCCTCACTGAGGCTTGTGCTGAGAGCTCAITGGGATGAA 1141
QY 278 nCysIleLeuMetLysGlnTyrlleGluPheValAlaAspArgLeuLeuValGluLeuGl 298
Db 1142 TTGCACCTCTAATGAAGCAATACATTTGAGTTTGTGGCAGACAGACTTATGCTGGAACTGG 1201
QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGl 318
Db 1202 TTTTAGCAGGTTTTCAGAGTAGAGAACCCATTTGACTTATGAGAGATATTTCACTGGA 1261
QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrglnArgPheAlaValMetAl 338
Db 1262 AGGAAAGACTAATCTTTTCAGAGAGAGTAGGCGAGTATCAGAGGATGGGAGTGCATGTC 1321
QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1322 AAGTCCACAGAGAATCTTTTACCTTGGATGCTGACTTC 1361
RESULT 15
US-10-843-641A-4196
; Sequence 4196, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals, Inc.
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; FILE OF INVENTION: Signature Gene Sets
; FILE REFERENCE: 689290-189
; CURRENT APPLICATION NUMBER: US/10/843,641A
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/873,367
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US/09/954,531
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,832
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/964,824
```



```
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/09/967,768
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/09/968,007
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,708
; PRIOR FILING DATE: 2001-10-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 8447
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4196
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-843-641A-4196

Alignment Scores:
Pred. No.: 4,04e-171 Length: 2500
Score: 1460.50 Matches: 277
Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
DB: 21 Gaps: 1

US-10-698-228-1 (1-351) x US-10-843-641A-4196 (1-2500)

Qy 4 ProGluArgProGluAlaAlaGlyLeuAspGluArgSerSerSerAspThrAs 23
Db |||||
Qy 302 CCTGAGCGGAGCCCGGCTCTGCGCAGCAAGACCGCGAGGAGGATCTTCCAGGAGCCAC 361
Db |||||
Qy 23 nGluSerGluLeuLysSer-----AsnGluGluProLeuLeuLysLeu 38
Db |||||
Qy 362 GGAGCGGAAACTAAGACGCTGCCCCGCGCGGAGATGAGCGCTGTGAGAGAAA 421
Db |||||
Qy 38 rSerArgArgPheValIlePhePheProIleGlnTyrProAspIleTyrLysMetTyrLysG 58
Db :|||||
Qy 422 CCCCCCGCGCTTGTTCATCTTCCCATCGAGTACCATGATATCTGCGCAGATGATATAAGAA 481
Db |||||
Qy 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysLeuProHisTr 78
Db :|||||
Qy 482 GGCAGAGGCTTCTTTTGGACCGCGGAGGAGTTGACCTCTCCAAGACATTCAGCACTG 541
Db |||||
Qy 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisTleuAlaPhePheAla 98
Db |||||
Qy 542 GGAATCCCTGAAACCCGAGGAGAGATATTTATATCCATGTTCTGGCTTCTTTGCGAGC 601
Db |||||
Qy 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
Db |||||
Qy 602 AAGCGATGGCATAGTAAATGAAACTTGGTGGAGCGATTTAGCCAGAGATTCAGATTAC 661
Db |||||
Qy 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
Db :|||||
Qy 662 AGAAGCCCGCTGTTTCTATGCTTCATAATGGCATGCAAAACATACATCTGAAATGTA 721
Db |||||
Qy 138 rSerLeuLeuLeuAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAenAl 158
Db :|||||
Qy 722 TAGCTCTCTTATTGACACTTACATAAAGATCCCAAGAAAGGGAATTTCTCTTCAATGC 781
Db |||||
Qy 158 aileGluThrMetProTyrValLysLysAlaAspTyrAlaLeuArgTyrIleAlaAs 178
Db |||||
Qy 782 CATTGAACGATGCTTGTCTCAAGAGAGAGGAGGAGCTGGGCTTGGCTGGATTGGGGA 841
Db |||||
Qy 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePh 198
Db :|||||
Qy 842 CAAAGAGGCTACCTATGTTGCAACGCTGTTGAGCTTTGCTGCTGAGTGAAGGCAATTTCTT 901
Db |||||
Qy 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
Db :|||||
Qy 902 TCCCGTCTCTTTTGGCTCGATATTTCTGGCTCAAGAAACGAGGACTGATGCTGCGCTCAC 961
Db |||||
Qy 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
Db |||||
```

Search completed: June 16, 2005, 18:52:15
Job time : 722 secs

```
Db 962 ATTTCTTAATGAACCTTATTAGCAGAGATGAGGTTTACCTGTGATTTTGCTTGCCTGAT 1021
Qy tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
Db |||||
Qy 1022 GTTCAAAACACCTGGTACACAAACCATCGGAGGAGAGTAAGAGAAATAATTATCAATGC 1081
Db |||||
Qy 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
Db |||||
Qy 1082 TGTTCGGATAGAACAGAGGATTTCTCCTCAGTGGCCTTGCCTGTGAAGCTCATTTGGGATGAA 1141
Db |||||
Qy 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuG 298
Db |||||
Qy 1142 TTGCACCTCTAATGAAGCAATACATTGAGTTTGTGGCAGACAGACTTATCTGGAACCTGGG 1201
Db |||||
Qy 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG 318
Db |||||
Qy 1202 TTTTAGCAGGTTTTCAGAGTAGAGAACCCATTTCACCTTATGGAGATATTTTCACTGGA 1261
Db |||||
Qy 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
Db |||||
Qy 1262 AGGAAAGACTAACTTCTTTGAGAGAGAGTAGGCGAGTATCAGAGGATGCGAGTGATGTC 1321
Db |||||
Qy 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db :|||||
Qy 1322 AAGTCCAACAGAGAAATTTCTTTTACCTTGGATGCTGACTTC 1361
Db |||||
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 16, 2005, 11:03:33 ; Search time 3856 Seconds
(without alignments)

3464.875 Million cell updates/sec

Title: US-10-698-228-1

Perfect score: 1821

Sequence: 1 MGDPERPEAGLDQDERSSS.....QRFVMAETTDNVFTLDADF 351

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Delop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10698228/runat_14062005_161416_28309/app_query.fasta_1.519
-DB=EST -QFMT=fastap -SUFFIX=rest -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10698228_@CNG_1_1_4352_@runat_14062005_161416_28309 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hgc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gse1:*
9: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1821	100.0	3292	CR617553	CR617553 full-leng
2	1821	100.0	4748	3 HSM802033	AL137348 Homo sapi
3	1791	98.4	4650	3 BC042948	BC042948 Homo sapi
4	1737.5	95.4	3501	3 CR627376	CR627376 Homo sapi
5	1591	87.4	1083	1 AL547501	AL547501 AL547501
6	1585	87.0	1056	9 AY398973	AY398973 Homo sapi
7	1468	80.6	1043	9 AY398975	AY398975 Mus muscu
8	1460.5	80.2	1571	3 CR603461	CR603461 full-leng
9	1460.5	80.2	1573	3 CR625489	CR625489 full-leng

10	1460.5	80.2	1582	3	CR608076	full-leng
11	1460.5	80.2	1588	3	CR602054	full-leng
12	1460.5	80.2	1592	3	CR621427	full-leng
13	1460.5	80.2	1600	3	CR604378	full-leng
14	1460.5	80.2	1605	3	CR590959	full-leng
15	1460.5	80.2	1612	3	CR614990	full-leng
16	1460.5	80.2	1613	3	CR609838	full-leng
17	1460.5	80.2	1623	3	CR618451	full-leng
18	1460.5	80.2	1630	3	CR602150	full-leng
19	1460.5	80.2	1796	3	CR596700	full-leng
20	1443	79.2	2088	3	AK088907	Mus muscu
21	1384	76.0	879	5	BQ441857	AGENCOURT
22	1354	74.4	892	7	CNI63214	952695 MA
23	1328.5	73.0	997	4	BM468712	AGENCOURT
24	1316	72.3	1008	9	AY398974	Pan trogl
25	1277	70.1	932	5	BUI96941	AGENCOURT
26	1275	70.0	914	7	CF995079	AGENCOURT
27	1257.5	69.1	946	7	CNO24619	AGENCOURT
28	1235	67.8	870	7	CO648380	ILLUMIGEN
29	1222	67.1	836	7	CR416843	CR416843
30	1221	67.1	893	5	BUI90680	AGENCOURT
31	1209.5	66.4	923	7	CR580780	CR580780
32	1209	66.4	1076	4	BM460735	AGENCOURT
33	1203	66.1	801	5	BP680185	BP680185
34	1197.5	65.8	933	6	CA983633	AGENCOURT
35	1195	65.6	1111	5	BM914217	AGENCOURT
36	1189	65.3	909	6	CA981614	AGENCOURT
37	1188	65.2	804	6	CD656661	AGENCOURT
38	1186.5	65.2	1028	5	BQ050629	AGENCOURT
39	1186	65.1	879	5	BUI89053	AGENCOURT
40	1186	65.1	910	6	CD794331	EST665692
41	1177	64.6	909	7	CK179258	EST768578
42	1170	64.3	928	4	BM018808	603646723
43	1168.5	64.2	1080	6	CD461172	EG09_04p0
44	1167	64.1	896	5	BUI70979	AGENCOURT
45	1164	63.9	838	5	BUI910710	AGENCOURT

ALIGNMENTS

RESULT 1	CR617553	3292 bp	mrna	linear	HTC 21-JUL-2004
LOCUS	full-length cDNA clone CSODI011YF14 of Placentia Cot 25-normalized				
DEFINITION	full-length cDNA clone CSODI011YF14 of Placentia Cot 25-normalized				
ACCESSION	CR617553				
VERSION	CR617553.1	GI:50498360			
KEYWORDS	HTC; CNSLT cDNA.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	Li, W.B., Gruber, C., Jesses, J. and Polayes, D.				
TITLE	Full-length cDNA libraries and normalization				
JOURNAL	Unpublished				
REMARK	Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue				
REFERENCE	2 (bases 1 to 3292)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)				
COMMENT	1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.				
FEATURES	Location/Qualifiers				
source	1..3292				
	/organism="Homo sapiens"				

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS01011yF14"
/tissue_type="placenta Cot 25-normalized"
/plasmid="pCWSPORT_6"

ORIGIN

Alignment Scores:

Pred. No.: 2,61e-210 Length: 3292
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-698-228-1 (1-351) x CR617553 (1-3292)

```

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
DB 27 ATGGGCGACCCGGAAGCGCGAAGCGCGCGGCTGGATCAGATCAGATCATCTTCA 86
QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
DB 87 GACACCAACGAAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 146
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
DB 147 CGGTTGTTCATCTTCCATCCAGTACCCTGATATTTGGAAATGTTATTAACAGGCACAG 206
QY 61 AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
DB 207 GCTTCCTCTGGACACAGAGAGGTCGATCTTATCAAGGATCTCCCTCACTGGAACAAG 266
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaLaserAsp 100
DB 267 CTTAAGCAGATGAGAAGTACTTCACTCTCACTCTAGCCCTTTTTCGAGCCAGTGAT 326
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 327 GGAATTTGAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTCGAGGTTCCAGAGGCT 386
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 387 CGCTGTTTCTATGGCTTTCAAAATCTCATCTCATCTTACCTTTCAGATGATGACGTTG 446
QY 141 LeuIleAspThrTyrIleArgPheProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 447 CTGATAGACACTTACATCAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 506
QY 161 ThrMetProTyrValLysLysLysAlaAspTyrAlaLeuArgTyrIleAlaAspArgLys 180
DB 507 ACCATGCCCTATGTTAAGAAAGAGAGATGGCGCTTCGATGGATAGCAGATAGAAA 566
QY 181 SerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePheSerGly 200
DB 567 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGGATTTCTTCTCAGGA 626
QY 201 SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 627 TCTTTTGGCTGCTATATCTTGGCTTAAAGAGAGAGGCTTATGTCAGGACTCACTTTTCC 686
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
DB 687 ATGAACTCATCAGCAGATGAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 746
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
DB 747 TACTTAGTAAATAAGCCTTCAGAAAGAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 806
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
DB 807 ATTGACAGGAGTATTTTAAACAGAGCCCTTCCAGTGTGGCCTCATTTGGAATGAAATTGCAT 866

```

```

QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
DB 867 TTGATGAACACAGTACATTGATTTGTAGCTGCAGAGATTACTTTGGAACTTTGGATCTTCA 926
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
DB 927 AAGTTTTTCAGCAGAGAAATCTTTTGAATTTATGGAAACATTTCTTTAGAGGAAAA 986
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
DB 987 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGCAGAAACC 1046
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB 1047 ACAGATAACGCTTTCACCTTGGATGCGAGATTTT 1079

RESULT 2
HSM802033
LOCUS Homo sapiens mRNA; cDNA DKFZp761E1312 (from clone DKFZp761E1312).
ACCESSION AL137348
VERSION AL137348.1 GI:6807859
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 4748)
AUTHORS Ansorge W., Krieger S., Regiert T., Rittmueller C., Schwager B.,
Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M. and
Wiemann S.
CONSTRM The German cDNA Consortium
TTITLE Direct Submission
JOURNAL Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFZp761E1312) is available at the RZPD Deutsches
Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp761E1312
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.

FEATURES             Location/Qualifiers
     source            1..4748
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="RZPD:DKFZp761E1312"
                     /db_xref="taxon:9606"
                     /clone="DKFZp761E1312"
                     /tissue_type="amygdala"
                     /clone_lib="761 (synonym: hamy2). Vector pSport1; host
                     DH10B; sites NotI + SalI"
                     /dev_stage="adult"
                     /note="ribonucleotide reductase"
     gene              1..4748
                     /gene="DKFZp761E1312"
                     /db_xref="GeneID:52545612"
                     /db_xref="GOA:Q9NTD8"
                     /translation="MGDPPEAAGLDQDSSSDTNESEIKSNEEPILLKSSRRFVI
                     FPQYDPDWKMYKQAQASFWTAESVDLSKDLPHNNKLKADESKYFISHILAFPAASDGI
                     VNENLVERFSQEVPEARCYFGQILLENHSEMYSLIDITYINDPKRFLFNALIE
                     TMYVKKKADWLRWIAIDRKSTGERVVAFAAEGVFVSGSFAAIFWLKKGLMPGLT
                     FSNELISRDEGLHCDFACLMFQVLVKNPSEERVEREIIIVDAVKIQEFLTEALPVGLIG

```

MNCILMKQYIEFVADRLIVELGFSKVFQAEINPDPFMEINISLEGKTNPEKRVSEYQRF
AVMAETTDNVFTILDADF"

ORIGIN

Alignment Scores:

Pred. No.: 4, 56e-210 Length: 4748
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-698-228-1 (1-351) x HSM802033 (1-4748)

Qy 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
Db 47 ATGGCGCACCAGAAAGCGCGGAGCGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 106
Qy 21 AspThrAsnGluSerGluIleLeuLysSerAsnGluProLeuLeuArgLysSerSerArg 40
Db 107 GACACCAACGAAAGTGAATGAATCAATCAAGAGCCACTCTCTAAGAAAGAGTCTCGC 166
Qy 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrPlysMetTyrLysGlnAlaGln 60
Db 167 CGGTTTGTCTATCTTTCATATCCAGTACCCCTGATATTTGGAAATGTATAACAGGCACAG 226
Qy 61 AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
Db 227 GCTTCCTTCTGGACAGAGAGGTGCTGATTTATCAAGAGATCTCCCTCACTGGAACAG 286
Qy 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAlaSerAsp 100
Db 287 CTTAAAGCAGATGAGAAGTACTTCTATCTCTCACATCTTAGCCTTTTGGCAGCGAGTAT 346
Qy 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 347 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGGCT 406
Qy 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 407 CGCTGTTTCTATGCTTTCATATCTCATGAGAAATGTTCACTCAGAGATGTACAGTTG 466
Qy 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 467 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 526
Qy 161 ThrMetProTyrValLysLysLysAlaAspTyrAlaLeuArgTyrIleAlaAspArgLys 180
Db 527 ACCATGCCCTATGTATTAAGAAAAAGCAGATGGCCCTTGCATGATGACAGATAGAAA 586
Qy 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
Db 587 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAAGAGATTTTCTTCTCAGGA 646
Qy 201 SerPheAlaAlaIlePheTyrPheLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 647 TCTTTTGTCTATTAATCTGTGCTAAAGAGAGAGGTCTTATGCCAGAGCTCATTCTTC 706
Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 707 AATGAATCATCAGCAGAGATGAAGACTTCTCACTGTGACTTTTGTCTGCTGATGTTCCAA 766
Qy 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
Db 767 TACTTAGTAAATAAGCCCTCAGAAAGAAAGGCTCAGGAGAGATCATTTGTGTGCTGTCAA 826
Qy 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 827 ATTGAGCAGAGGTTTTTAACAGAGAGCCCTTCCAGTGTGGCTCATTTGAATGAATGCATT 886
Qy 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 887 TTGATGAAACAGTACATTGATTGTTAGCTGACAGATTACTTTGTGGAACCTTGGATTCTCA 946

Qy 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 947 AAGGTTTTTTCAGGCAGAAAAATCCTTTTGATTTTATGAAAAATCTTTTAGAAGAAAA 1006
Qy 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 1007 ACAAATTTCTTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTATATGCGAGAAACC 1066
Qy 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1067 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1099

RESULT 3

BC042948 4650 bp mRNA linear HTC 12-OCT-2004
LOCUS BC042948.2 GI:34194000
DEFINITION Homo sapiens cDNA clone IMAGE:4798175, containing frame-shift errors.

ACCESSION BC042948

VERSION BC042948.2

KEYWORDS HTC.

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

1 (bases 1 to 4650)

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,

Klausner, R.D., Collins, P.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,

Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,

Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,

Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,

Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,

Sheehez, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,

Carninci, P., Brange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,

Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,

Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,

Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,

Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,

Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,

Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,

Butcher, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 4650)

Director MGC Project.

Direct Submission

Submitted (09-JAN-2003) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.ncbi.nih.gov>

On Aug 25, 2003 this sequence version replaced gi:27695575.

Contact: MGC help desk

Email: cgabs-remail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki

Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Genome Sequence Centre,

BC Cancer Agency, Vancouver, BC, Canada

info@bcgsc.bc.ca

Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,

Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth

Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,

Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,

Teika Odonal, Diana Maquie, Anca Petrescu, Anna Liisa Prabh,

Parvaneh Saeedi, JR Santos, Angelique Schnerch, Ursula Skalska,

Duane Smalus, Jeff Stott, Miranda Tsai, George Yang, Jacque

Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 70 Row: h Column: 19
This clone has the following problem: frame shifted.

FEATURES

source

Location/Qualifiers
1. .4650
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4798175"
/issue_type="Brain, hypothalamus"
/clone_lib="NIH MGC_96"
/lab_host="DH10B"
/note="Vector: pBluescript"

ORIGIN

Alignment Scores:

Pred. No.: 2,04e-206 Length: 4650
Score: 1791.00 Matches: 350
Percent Similarity: 99.72% Conservative: 0
Best Local Similarity: 99.72% Mismatches: 1
Query Match: 98.35% Indels: 1
DB: 3 Gaps: 0

US-10-698-228-1 (1-351) x BC042948 (1-4650)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
DB 72 ATGGCGGACCGGAAGCGCGGCTGGATCAGATGAGATCATCTTCA 131
QY 21 AspThrAsnGluSerGluLeuLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
DB 132 GACACCAAGCAAGTAAAGTAAATCAATGAAGACA-CTCTAAGAAAGAGTTCTGC 190
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
DB 191 CGGTTTGTCATCTTCCATCCAGTACCCTGATATTGGAAATGTATAAACAGGCACAG 250
QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
DB 251 GCTTCTCTTGACAGCAGAGAGGTCGACTTATCAAGGATCTCCCTCACTGGAACAAG 310
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
DB 311 CTAAAGCAGATGAGAAGTACTTCACTCTCACTTAGCCCTTTTTCAGCCAGTGAT 370
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 371 GGAATGTAAATGAATAATTTGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 430
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 431 CGCTGTTTCTATGGCTTTCAAAATTCATCAGAAATGTTCACTCAGAGATGTACAGTTG 490
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 491 CTGATAGACACTTACATCAGATCCCAAGAAAGGGAATTTTATTATGCAATGTAA 550
QY 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
DB 551 ACATGCCCTTATGTTTAAAGAAAGACAGATTTGGCCCTTGGCATGGATGAGATAGAAA 610
QY 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
DB 611 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 670
QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 671 TCTTTTGTGCTATATCTGGCTAAAGAGAGAGAGTCTTATGCCAGGACTCACTTTTCC 730
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240

DB 731 AATGNACTCATCAGCAGATGAAGGACTTCACTGTGACTTTGCTTGCCTCATGTTCCAA 790
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
DB 791 TACTTAGTAATAAGCCTTCAGAAAGGGTCAGGAGATCATTTGTGTGCTGTCAA 850
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
DB 851 ATTGAGCAGGAGTTTAAACAGAAAGCCTTCCAGTTGGCCTCATTTGGAATGAATGCATT 910
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
DB 911 TTGATGAACAGTACATTCAGTTTGTAGCTGCACAGATTACTTTGGAACTTGGATTTCA 970
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
DB 971 AAGTTTTTCAGCAGAAATCCTTTGATTTTATGGAACATTTCTTTAGAGAGAAA 1030
QY 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
DB 1031 ACAATTTCTTTTCAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC 1090
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB 1091 ACAGATAACGTCCTTCACCTTTGGATGCAGATTTT 1123

RESULT 4
CR627376
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
CONSTRM
TITLE
JOURNAL
COMMENT
FEATURES
source
gene
CDS

CR627376 3501 bp mRNA linear HTC 22-SEP-2004
Homo sapiens mRNA; cDNA DKFZp686M05248 (from clone DKFZp686M05248).
CR627376
CR627376.1 GI:50949847
HTC.
Homo sapiens (human)
Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3501)
Ortenwaelder, B., Obermaier, B., Deutschenbaier, S., Schaipp, A.,
Mewes, H.W., Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and
Wiemann, S.
The German cDNA Consortium
Direct Submission
Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Medigenomix (Martinried/Germany) within the cDNA
sequencing consortium of the German Genome Project.
This clone (DKFZp686M05248) is available at the RZPD Deutsches
Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
<http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp686M05248>
Further information about the clone and the sequencing project is
available at <http://mips.gsf.de/projects/cdna/>.
Location/Qualifiers
1. .3501
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686M05248"
/tissue_type="salivary gland"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
/note="ribonucleotide reductase M2 B (TP53 inducible)"
1. .3501
/gene="DKFZp686M05248"
332. .1225
/gene="DKFZp686M05248"
/codon_start=1
/product="hypothetical protein"

/protein_id="CAH10473.1"
 /db_xref="GI:50949848"
 /translation="MYKQASFWTTAEVDLSKDLPHNNKKADEKYFISHILAFPA
 SDGIVNENLVERFSQVQVPEARCFYGFQILINVENHSEMYSLIIDTYIRDPKKREFLF
 NATETWPKVKKADWALRWADKSTFGERVAFAAVEGVFFSGSPAIIFWLKKRGLM
 PGLTFNELLSDREGLHCDPACLMFOVLNKPPEERVRELIIVDAVKLEQSLFEALPV
 GLIGMCLMKQYIEFFVADRLVVELGFSKVFQAEINPDPFENISLEKNTFFEKRVSE
 YQRFVNAETTDNVTLDADF"

ORIGIN

Alignment Scores:
 Pred. No.: 4 5e-200 Length: 3501
 Score: 1737.50 Matches: 339
 Percent Similarity: 96.06% Conservatives: 2
 Best Local Similarity: 95.49% Mismatches: 7
 Query Match: 95.41% Indels: 7
 DB: 3 Gaps: 1

US-10-698-228-1 (1-351) x CR627376 (1-3501)

QY 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu----- 16
 |||||
 Db 158 CCGCTCGCTGCTGCGCGGGCCAGCAGCAGCCTTGGAGCCTTCAGCCGGG 217
 |||||

QY 17 ArgSerSerAspThrAsnGluSerGluLeuLysSerAsnGluProLeuLeuArg 36
 |||||
 Db 218 AGATCATCTTCAGACACCAACGAAAGTGAATAAGTCAATGAAGACCACTCCTAAGA 277
 |||||

QY 37 LysSerSerArgArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyr 56
 |||||
 Db 278 AAGAGTTCTCGCGGTTGTTCATCTTCCCAATCAGTACCCTGATATTTGGAAATGTAT 337
 |||||

QY 57 LysGlnAlaGlnAlaSerPheTyrAlaGluValAspLeuSerLysAspLeuPro 76
 |||||
 Db 338 AATAGCAGCAGCGCTTCTTCTGACAGCAGAGGTCGACTTATCAAGAGATCTCCCT 397
 |||||

QY 77 HisTrpAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePhe 96
 |||||
 Db 398 CACTGGAACAGCTTAAAGCAGATGAGAAGTACTTCTCATCTCTCACATCTTAGCCTTTT 457
 |||||

QY 97 AlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGln 116
 |||||
 Db 458 GCAGCAGTGATGAAATGTAAATGAAAAATTTGGTGAACGCTTTAGTCAGGAGGTCCAG 517
 |||||

QY 117 ValProGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGlu 136
 |||||
 Db 518 GTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAAATCTCATCAGAGATGTTCACTCAGAG 577
 |||||

QY 137 MetTyrSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPhe 156
 |||||
 Db 578 ATGTACAGTTTGTGTAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTT 637
 |||||

QY 157 AsnAlaIleGluThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTyrIle 176
 |||||
 Db 638 AATGCAATTGAAACCATGCTTATTTAGAAAAAAGCAGATTTGGGCTTGGCATGGATA 697
 |||||

QY 177 AlaAspArgLysSerThrPheGlyGluArgValValAlaPheAlaAlaValGluVal 196
 |||||
 Db 698 GCAGATAGAAAAATCTACTTTTGGGGAAGAGTGGTGGCCCTTGTCTGTAGAGGAGTT 757
 |||||

QY 197 PhePheSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGly 216
 |||||
 Db 758 TTCTTCTCAGGATCTTTTGTGCTATATTTCTGCTTAAAGAGAGAGGTCTTATGCCAGGA 817
 |||||

QY 217 LeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCys 236
 |||||
 Db 818 CTCACATTTTTCATGAATCACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGC 877
 |||||

QY 237 LeuMetPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleVal 256
 |||||
 Db 878 CTGATGTTCCAACTACTAGTAAATAAGCTTCCAGAGAAAGGGTCCAGGAGATCATTTGT 937
 |||||

QY 257 AspAlaValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGly 276
 |||||

Db 938 GATGCTGTCAAAATTTGACAGAGGATTTTAAACAGAGCCTTCCAGTTGGCCTCATTTGA 997
 |||||

QY 277 MetAsnCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGlu 296
 |||||

Db 998 ATGAATTTGATTTTATGATGAACAGATACATTGAGTTTGTAGCTGACAGATTACTTTGGAA 1057
 |||||

QY 297 LeuGlyPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSer 316
 |||||

Db 1058 CTTGGATTTCTCAAAGGTTTTTTCAGCAGAAATCTTTTATGATTTATGAAAAATTTCT 1117
 |||||

QY 317 LeuGluGlyLysThrAsnPhePheGlyLysArgValSerGluTyrGlnArgPheAlaVal 336
 |||||

Db 1118 TTAGAAGGAAACAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTT 1177
 |||||

QY 337 MetAlaGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
 |||||

Db 1178 ATGCAGAAACACAGATAACGCTTTCACCTTGGATGAGATTTT 1222
 |||||

RESULT 5
 AL547501 1083 bp mRNA linear EST 25-MAR-2004
 LOCUS
 DEFINITION AL547501 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 clone CSODI011YF14 5-PRIME, mRNA sequence.
 ACCESSION AL547501
 VERSION AL547501.3 GI:45747948
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1083)
 AUTHORS Li, W.-B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT On Feb 15, 2001 this sequence version replaced gi:31269332.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 4436.r
 For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?s=CSODI011DC07QPI&c=4436.r.

FEATURES

source
 1. 1083
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clones="CSODI011YF14"
 /tissue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Alignment Scores:
 Pred. No.: 5 81e-183 Length: 1083
 Score: 1591.00 Matches: 319
 Percent Similarity: 97.85% Conservatives: 0
 Best Local Similarity: 97.85% Mismatches: 5
 Query Match: 87.37% Indels: 4
 DB: 1 Gaps: 0

US-10-698-228-1 (1-351) x AL547501 (1-1083)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSe 20

Db	27	ATGGGCGACCCGGAAGCGCGCGGCTGGATCAGATGAGAGATCATCTTC	86
Qy	20	rAspThrAsnGluSerGluLeuLeuSerAsnGluGluProLeuLeuArgLysSerSerAr	40
Db	87	AGACACCAACGAAGTGAATTAAGTCAATGAAGGCCACTCTTAAGAAAGAGTCTCG	146
Qy	40	gArgPheValIlePheProIleGlnTyrProAspIleTyrPheLysMetTyrLysGlnAlaGl	60
Db	147	CCGGTTTGTTCATCTTCCCAATCAGTACCTGATATTGGAAAATGTATAAACAGGCACA	206
Qy	60	nAlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrPheSerLys	80
Db	207	GGCTTCCTTCGGACAGCAGAGAGTGCATTTATCAAAAGGATCTCCCTCAGTGGACAA	266
Qy	80	sLeuLysAlaAspGluLeuLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAs	100
Db	267	GCTTAAGCAGATGAGAGTACTTTCATCTCTACATCTTAGCCTTTTTCAGCCAGTGA	326
Qy	100	pGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAl	120
Db	327	TGGAATTGTAATGAAATTTGGTGAGCGCTTAGTCAGGAGTGCAGGTTCCAGAGGC	386
Qy	120	aArgCysPheTyrGlyPheGlnIleLeuLeuGluAsnValHisSerGluMetTyrSerLe	140
Db	387	TCGCTGTTCTATGGCTTTCAAAATCTCATCGAATGTTCACTCAGAGATGTACAGTTT	446
Qy	140	uLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGl	160
Db	447	GCTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATCAATTGA	506
Qy	160	uThrMetProTyrValLysLysLysAlaAspTyrAlaLeuArgTyrPheAlaAspArgLy	180
Db	507	ACCATGCCCTATGTAAGAAAAGCAGATTTGGGCTTGGCATGATAGCAGATAGAA	566
Qy	180	sSerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGl	200
Db	567	ATCTACTTTTGGGAAAGAGTGGTGGCTTGTGCTGTAGAGGAGTTTCTTCTCAGG	626
Qy	200	ySerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSe	220
Db	627	ATCTTTTGTCTATATTCTGGCTAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC	686
Qy	220	rAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGl	240
Db	687	CAATGAACTCNCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCTGATGTCCA	746
Qy	240	nTyrLeuValAsnLysProSerGluArgValArgGluIleIleValAspAlaVally	260
Db	747	ATACTTAGTAAATAAGCCTTCAGAAAGAGGTCCAGGAGATCATTTGTTGATGCTGCA	806
Qy	260	sIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIl	280
Db	807	AATTGAGCAGAGTGTTTTAAACAGAGCCTTGCAGTTGGCCTCATTTGGAATGAATGCAT	866
Qy	280	eLeuMetLysGlnTyrIleGluPheValAlaAspArgLeu-LeuValGluLeuGlyPheS	300
Db	867	TTTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAMCTTGATCT	926
Qy	300	erLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyL	320
Db	927	CAAGGKTTTCAGGCAGAAAATCCCTTTGATTATGGAAM-AITTTCTTTAGAGGAA	985
Qy	320	yThrAsnPhePhe 324	
Db	986	AAA-AAATTTCTTT 998	

RESULT 6

AY398973
LOCUS Homo sapiens HCM0069 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY398973

VERSION	AY398973.1	GI:39754962
KEYWORDS	GSS.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS	1 (bases 1 to 1056) Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.	
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios	
JOURNAL	Science 302 (5652), 1960-1963 (2003)	
PUBLISHED	14671302	
REFERENCE	2 (bases 1 to 1056) Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.	
AUTHORS	Direct Submission	
TITLE	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA	
JOURNAL	This sequence was made by sequencing genomic exons and ordering them based on alignment.	
COMMENT	Location/Qualifiers	
FEATURES	1..1056 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" <1..>1056 /locus_tag="HCM0069"	
source		
gene		
ORIGIN		
Alignment Scores:		
Pred. No.:	3.02e-182	Length: 1056
Score:	1585.00	Matches: 313
Percent Similarity:	89.17%	Conservative: 0
Best Local Similarity:	89.17%	Mismatches: 38
Query Match:	87.04%	Indels: 0
DB:	9	Gaps: 0
US-10-698-228-1 (1-351) x AY398973 (1-1056)		
Qy	1	MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20
Db	1	ATGGCGACCCGGAAGCGCGGATGAGATGAGATGAGATCATCTTCA 60
Qy	21	AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerSerArg 40
Db	61	GACACCAACGAAAGTGAATTAAGTCAATGAAGGCCACTCTCTAAGAAAGAGTTCTCGC 120
Qy	41	ArgPheValIlePheProIleGlnTyrProAspIleTyrPheLysGlnAlaGln 60
Db	121	CGGTTTGTCTATCTTCCAAATCCAGTACCTGATATTGGAAAAATGTATAACAGGCACAG 180
Qy	61	AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrPheAsnLys 80
Db	181	GCTTCCTCTTGGACAGCAGAGAGTGCATTTATCAAGGATCTCCCTCAGTGGACCAAG 240
Qy	81	LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Db	241	CTTAAAGCAGATGAGAAGTACTTCTCATCTCTACATCTTAGCCTTTTTCAGGCCAGTGTAT 300
Qy	101	GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db	301	GGAAATGTAATGAAAAATTTGGTGAGCGCTTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
Qy	121	ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db	361	CGCTGTTTCTATGGCTTTCAAAATTTCTATCAGAGATGTTTCACTCAGAGATGTACAGTTTG 420
Qy	141	LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160


```

|||||
781 ATCGAGCA-----GAAGCCTTGGCTGTGGCTCATTGGATGAATGTGT 827
|||||
281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuGlyPheSer 300
|||||
828 TTGATGAACAGTACATGGTTGTAGCTGACAGATTACTTGGAGAACTTGGATTCTCA 887
|||||
301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
|||||
888 AAGATTTTTCAGCGAGAAAATCCCTTGTATTTATGAGAACATTTCAATTAGAGGAAA 947
|||||
321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
|||||
948 ACAAAATTTCTTTGAGAAACGAGTTTCTGAGTATCAGCGATTTCGATCGTCCGAAACC 1007
|||||
341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
|||||
1008 ACAGATAATGTCTTACCTTGGATGCGATTTC 1040

CR603461 1571 bp mRNA linear HTC 21-JUL-2004
full-length cDNA clone CS0DH002YP23 of T cells (Jurkat cell line)
of Homo sapiens (human).
ACCESSION CR603461
VERSION CR603461.1 GI:50484268
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1571)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 1571)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
source 1..1571
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DH002YP23"
/tissue_type="T cells (Jurkat cell line)"
/plasmid="pCMVSPORT_6"

ORIGIN
Alignment Scores:
Pred. No.: 8..86e-167 Length: 1571
Score: 1460..50 Matches: 277
Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
DB: 3 Gaps: 1

US-10-698-228-1 (1-351) x CR603461 (1-1571)

QY 4 ProGluArgProGluAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
|||||
Db 144 CCTGAGCGGACCCGGCTCTGCGCCAGCAAGACCGGAGGAGGATCTTCCAGGAGGCCAC 203
|||||

```

```

QY 23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
|||||
Db 204 GGAGCGGAACCTAAAGAGAGCTGCCCGCGCGTGGAGGATGAGCGCTGCTGAGAGAAA 263
|||||
QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysG1 58
|||||
Db 264 CCCCCCGCGCTTGTCTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAGAA 323
|||||
QY 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysLeuProHisTr 78
|||||
Db 324 GGCAGAGGCTTCTCTTTGGACCGCGGAGGAGGTGGACCTCTCCAAGGACATTCAGCACTG 383
|||||
QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaA 98
|||||
Db 384 GGAATCCCTGAAACCCGAGGAGAGATATTTATATCCCATGTCTTGGCTTCTTCTTGCAGC 443
|||||
QY 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
|||||
Db 444 AGCGATGGCATAGTAATGAAACTTGTGTGGAGGATTTAGCCAGAGTTTCAGATTAC 503
|||||
QY 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
|||||
Db 504 AGAAGCGCGCTGTCTTCTATGCTTCCAAATTTGCCATGGAACATACATCTGAAATGTA 563
|||||
QY 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
|||||
Db 564 TAGTCTTCTTATTGACACTTACATATAAAGATCCCAAGAAAGGGAATTTCTTCTCAATGC 623
|||||
QY 158 alleGluThrMetProTyrValLysLysAlaAspTyrAlaLeuArgTyrPheAlaAs 178
|||||
Db 624 CATGAAACGATGCCCTTGTGTCAAGAAGAGGAGACTGGCCCTTGGCTGGATGGGA 683
|||||
QY 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePh 198
|||||
Db 684 CAAAGAGGCTACTATGTGTAACGTGTGTAGCTTTGCTGCTGAGGAGCAATTTCTTCT 743
|||||
QY 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
|||||
Db 744 TTCGGTCTTTTTCGGTCGATATTTCTGGCTCAAGAAACGAGGACTGTATGCTGCCTCAC 803
|||||
QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
|||||
Db 804 ATTTTCTAATGAACCTTATTAGCAGAGATGAGGGTTTACACTGTGATTTTGTGCTGAT 863
|||||
QY 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
|||||
Db 864 GTTCAACACCTGGTACACAAACCTCGAGGAGAGTAAAGAGAAATAATATCAATGC 923
|||||
QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
|||||
Db 924 TGTTCCGATAGAACAGGAGTTCTCTCACTGAGGCTTCTGCTGTAAGCTCATTTGGGATGAA 983
|||||
QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluG1 298
|||||
Db 984 TTGCACCTTAATGAACCAATACATTTGAGTTTGTGGCAGACAGACTTATGCTGGAACTGGG 1043
|||||
QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG1 318
|||||
Db 1044 TTTTAGCAAGGTTTTCAGATGAGAACCCCATTTTCACTTTATGAGAGATATTTCACTGGA 1103
|||||
QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
|||||
Db 1104 AGGAAAGACTTAACCTTCTTTGAGAAGAGAGTAGGGGAGTATCAGAGGATGGGAGTGATGTC 1163
|||||
QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
|||||
Db 1164 AAGTCCCAACAGAAATTTCTTTTACCTTGGATGCTGACTTC 1203
|||||

RESULT 9
CR625489 1573 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CS0DA003YL09 of Neuroblastoma of Homo
DEFINITION sapiens (human).

```

```

ACCESSION   CR625489
VERSION     CR625489.1 GI:50506296
KEYWORDS    HTC; CNSLUT_CDNA.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1573)
AUTHORS     Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE       Full-length cDNA libraries and normalization
JOURNAL     Unpublished
REMARK      Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
            Faraday Avenue
            2 (bases 1 to 1573)
REFERENCE   2 (bases 1 to 1573)
AUTHORS     Genoscope.
TITLE       Direct Submission
JOURNAL     Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
FEATURES    Location/Qualifiers
            1..1573
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="CS0DA003YL09"
             /issue_type="Neuroblastoma"
             /plasmid="pCMVSPORT_6"
ORIGIN
Alignment Scores:
Pred. No.:      8,87e-167      Length:      1573
Score:          1460.50        Matches:    277
Percent Similarity: 88.42%      Conservative: 36
Best Local Similarity: 78.25%    Mismatches: 35
Query Match:     80.20%         Indels:     6
DB:              3             Gaps:       1

US-10-698-228-1 (1-351) x CR625489 (1-1573)

QY      4 ProGluArgProGluAlaGlyLeuAspGlnAspGlu-ArgSerSerAspThrAs 23
D      155 CTGAGCGGAGCCCGCGTCTCGCCAGCAAGACCGGAGGAGGATCTCCAGGAGCCAC 214
QY      23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuLysSe 38
D      215 GGAGCGGAAACTAAGACGAGTCCCGCGGCGTGGAGGATGAGCCCTGCTGAGAGAAA 274
QY      38 rSerArgArgPheValIlePheProIleGlnTyProAspIleTrpLysMetTyLysG 58
D      275 CCCCAGCGCTTGTTCATCTTCCCATCGAGTACCATGATATCTGCAGATGTATAAGAA 334
QY      58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
D      335 GGCAGAGGCTTCTTTTGGACCGCCGAGGAGGTGGACCTCTCCAGAGCATTCAGCACTG 394
QY      78 pAsnLysLysLysAlaAspGluLysTyPheIleSerHisLleLeuAlaPheAlaAl 98
D      395 GGAATCCCTGAAACCCGAGAGAGATATTTATATCCCATGTTCTGGCTTTCTTCGAGC 454
QY      98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
D      455 AAGCGATGCGATAGTAATGAAACTTGTGTGGAGCATTTAGCCAGAGAGTTCAGATTAC 514
QY      118 oGluAlaArgCysPheTyrglyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
D      515 AGAAGCCCGCTGTTTCTATGGCTTCCAAATTTGCCATGGAAACATACATCTTCAAATGTA 574

```

138 rSerLeuLeuIleAspThrTyrlleArgAspProIlyLysArgGluPheLeuPheAsnAl 158
575 TAGTCTTCTTATTGACACTTACATAAAGATCCCAAGAAAGGGGAATTTCTCTTCAATGC 634
QY 158 aileGluThrMetProTyrvallLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
635 CATTGAACCATGCTGTGTCAAGAAAGGAGAGAGCTGGGCTTGGCTGGATGGGGA 694
QY 178 pArgLysSerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePh 198
695 CAAAGAGGCTACCTATGTGTGAACGCTTGTAGCTTGTCTCAGTGAAGGCAATTTCTT 754
QY 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
755 TTCGGTCTTTTTCGTCGATATTCTGGCTCAAGAAACGAGGACTGATGCTGGGCTCAC 814
QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
815 ATTTCTTAATGAACCTTATTAGCAGATGAGGGTTTACACTGTGTAATTTGCTGCTGAT 874
QY 238 tPheGlnTyLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
875 GTTCAACACCTGTCACAAACCATCGAGGAGAGAGTAAGAGAAATATTATCAATGC 934
QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
935 TGTTCGATAGAACAGGAGTCTCTCACTGAGGCTTGTGCTGTGAAGCTCATTTGGGATGAA 994
QY 278 nCysIleLeuMetLysGlnTyrlleGluPheValAlaAspArgLeuValGluLeuGl 298
995 TTGCACCTTAATGAAGCAATACATTTGAGTTTGTGCGACAGACACTTATGTGGAACTGGG 1054
QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGl 318
1055 TTTTAGCAGGTTTTCAGATAGAGAACCCATTTGACTTTATGGAGAAATATTCTACTGA 1114
QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrglnArgPheAlaValMetAl 338
1115 AGGAAGACTTAACCTCTTTCAGAAAGAGAGTAGGCGAGTATCAGAGGATGGGATGATGTC 1174
QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
1175 AAGTCCCAACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1214

RESULT*10
CR608076
LOCUS
DEFINITION CR608076 1582 bp mRNA linear HTC 21-JUL-2004
full-length cDNA clone CS0DK012YD23 of HeLa cells Cot 25-normalized
of Homo sapiens (human).
ACCESSION CR608076
VERSION CR608076.1 GI:50488883
KEYWORDS HTC; CNSLUT_CDNA.
SOURCE Homo sapiens
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1582)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue
 2 (bases 1 to 1582)
REFERENCE Genoscope.
AUTHORS Direct Submission
TITLE Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a

Qy	258	aVallyslleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs	278
Db	922	TGTTCCGATAGAACAGAGAGTCTCTCACTAGGCGCTTCCTGTGAAGCTCATTTGGGATGAA	981
Qy	278	nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuG1	298
Db	982	TTGCACCTCTATGAGCAATACATTGAGTTTGTGGCAGACAGACTTATGCTGGAACTGGG	1041
Qy	298	yPheSerLysValPheGlnAlaGluAanPropheAspPheMetGluAanIleSerLeuG1	318
Db	1042	TTTTTAGCAAGTTTTCAGAGTAGAGAACCCATTGACTTTATGTCGAGAATAATTTCACTCGA	1101
Qy	318	uGlyLysThrAsnPhesPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl	338
Db	1102	AGGAAAGACTTAACCTTCTTTCAGAGAGAGTAGCGGAGTATCAGAGGATGGGAGTGATGTC	1161
Qy	338	aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe	351
Db	1162	AAGTCACACAGAGAAATCTTTTACCTTGGAGCTGACTTC	1201
RESULT 11			
CR602054			
LOCUS			
DEFINITION		full-length cDNA clone CSODA011VH14 of Neuroblastoma Of Homo sapiens (human).	
ACCESSION		CR602054	
VERSION		CR602054.1	GI:50482861
KEYWORDS		HTC; CNSLT cDNA.	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS		Li, W.B., Gruber, C., Jessee, J. and Polayes, D.	
TITLE		Full-length cDNA libraries and normalization	
JOURNAL		Unpublished	
REMARK		Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Paraday Avenue	
REFERENCE		2 (bases 1 to 1588)	
AUTHORS		Genoscope.	
TITLE		Direct Submission	
JOURNAL		Submitted (20-JUL-2004) Genoscope - Centre National de Sequençage	
COMMENT		BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr) - Web : www.genoscope.cns.fr)	
FEATURES		1st strand cDNA was primed with a NotI-oligo(GT) primer. Five primer end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.	
source		Location/Qualifiers	
		1..1588	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="CSODA011VH14"	
		/tissue_type="Neuroblastoma"	
		/plasmid="pCMVSPORT_6"	
ORIGIN			
Alignment Scores:			
Pred. No.:	9e-167	Length:	1588
Score:	1460.50	Matches:	277
Percent Similarity:	88.42%	Conservative:	36
Best Local Similarity:	78.25%	Mismatches:	35
Query Match:	80.20%	Indels:	6
DB:	3	Gaps:	1
US-10-698-228-1 (1-351) x CR602054 (1-1588)			
Qy	4	ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerAspThrAs	23
Db	142	CCTGAGCGGACCGCGCTCTGGCCAGCAGACCGCAGGAGGATCTTCAGAGGCCAC	201

Qy	258	aVallyslleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs	278
Db	922	TGTTCCGATAGAACAGAGAGTCTCTCACTAGGCGCTTCCTGTGAAGCTCATTTGGGATGAA	981
Qy	278	nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuG1	298
Db	982	TTGCACCTCTATGAGCAATACATTGAGTTTGTGGCAGACAGACTTATGCTGGAACTGGG	1041
Qy	298	yPheSerLysValPheGlnAlaGluAanPropheAspPheMetGluAanIleSerLeuG1	318
Db	1042	TTTTTAGCAAGTTTTCAGAGTAGAGAACCCATTGACTTTATGTCGAGAATAATTTCACTCGA	1101
Qy	318	uGlyLysThrAsnPhesPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl	338
Db	1102	AGGAAAGACTTAACCTTCTTTCAGAGAGAGTAGCGGAGTATCAGAGGATGGGAGTGATGTC	1161
Qy	338	aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe	351
Db	1162	AAGTCACACAGAGAAATCTTTTACCTTGGAGCTGACTTC	1201
RESULT 11			
CR602054			
LOCUS			
DEFINITION		full-length cDNA clone CSODA011VH14 of Neuroblastoma Of Homo sapiens (human).	
ACCESSION		CR602054	
VERSION		CR602054.1	GI:50482861
KEYWORDS		HTC; CNSLT cDNA.	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS		Li, W.B., Gruber, C., Jessee, J. and Polayes, D.	
TITLE		Full-length cDNA libraries and normalization	
JOURNAL		Unpublished	
REMARK		Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Paraday Avenue	
REFERENCE		2 (bases 1 to 1588)	
AUTHORS		Genoscope.	
TITLE		Direct Submission	
JOURNAL		Submitted (20-JUL-2004) Genoscope - Centre National de Sequençage	
COMMENT		BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr) - Web : www.genoscope.cns.fr)	
FEATURES		1st strand cDNA was primed with a NotI-oligo(GT) primer. Five primer end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.	
source		Location/Qualifiers	
		1..1588	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="CSODA011VH14"	
		/tissue_type="Neuroblastoma"	
		/plasmid="pCMVSPORT_6"	
ORIGIN			
Alignment Scores:			
Pred. No.:	9e-167	Length:	1588
Score:	1460.50	Matches:	277
Percent Similarity:	88.42%	Conservative:	36
Best Local Similarity:	78.25%	Mismatches:	35
Query Match:	80.20%	Indels:	6
DB:	3	Gaps:	1
US-10-698-228-1 (1-351) x CR602054 (1-1588)			
Qy	4	ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerAspThrAs	23
Db	142	CCTGAGCGGACCGCGCTCTGGCCAGCAGACCGCAGGAGGATCTTCAGAGGCCAC	201

```

QY 23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
Db 202 GGAGCGCAAACTAAAGCAGCTGCCCGCGCGTGGAGGATGAGCGCTGCTGAGAGAAA 261
QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysG1 58
Db 262 CCCCCCGCGCTTTGTCTATCTCCCATCGAGTATCCATGATATCTGGCAGATGATATAGAA 321
QY 58 nAlaGlnAlaSerPheTyrPheGlnIleLeuValAspLeuSerLysAspLeuProHisTr 78
Db 322 GGAGAGGCTTCTTTTGGACCGCGAGGAGGTGGACCTCTCCAGGACATTCAGCACTG 381
QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAl 98
Db 382 GGAATCCCTGAAACCGCGAGGAGATATTTATATCCATGCTTCTGCTTCTTTGCGAGC 441
QY 98 aSerAspGlyIleValAsnGlnAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
Db 442 AAGCATGGCATAGTAATAAGAACTTGGTGGAGCGATTTAGCCAAAGTTTCAGATTAC 501
QY 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
Db 502 AGAAGCCCGCTGTTCTATGCTTCCAAATTGCATGGAAACATACATCTCGAAATGTA 561
QY 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
Db 562 TAGTCTCTATTGACACTTACATAAAGATCCCAAGAAAGGGAATTTCTCTCANTGC 621
QY 158 alleGluThrMetProTyrValLysLysLysAlaAspTyrPAlaLeuArgTrpIleAlaAs 178
Db 622 CATTGAAACGATGCTTGTCTCAAGAAGAGGAGGAGCTGGCGCTTGGCTGGATTTGGGA 681
QY 178 pArgLysSerThrPheGlyGluArgValAlaAlaPheAlaAlaValGluGlyValPhePh 198
Db 682 CAAAGAGGCTACCTATGTGTGAACGTGTGTAGCTTTTGTGCTGCAAGTGAAGGCAATTTCTT 741
QY 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
Db 742 TTCGGTCTTTTGGTGGATATCTGCTCAAGAAAGAGGAGTGTATGCTGGCTGCCTCAC 801
QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
Db 802 ATTTCTAATGAATTTATAGCAGATGAGGTTTACACTGTGATTTGCTTGCCTGAT 861
QY 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
Db 862 GTTCAAAACACCTGTTACAAACCATCGGAGGAGAGATGAAGAGAAATAATATCAATGC 921
QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValClyLeuIleGlyMetas 278
Db 922 TGTTCGGATAGAACAGAGTCTCTCACTGAGGCTTGCCTGTGAAGCTCATTTGGATGAA 981
QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuG1 298
Db 982 TTGCATCTTAATGAGCAATACATTGATTTGGTCAGACAGACTTATGCTGGACTGGG 1041
QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG1 318
Db 1042 TTTTAGCAAGTTTTCAGAGTAGAAGAACCATTTGACTTTATGAGAGATATTTCACTGGA 1101
QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
Db 1102 AGAAAGACTTAATCTCTTTGAGAGAGAGTAGGCGGAGTATCAGAGGATGGAGTGTATGTC 1161
QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1162 AGTCCCAACAGAAATCTTTTACCTTGGATGCTGACTTC 1201

```

RESULT 12

CR621427

LOCUS

DEFINITION full-length cDNA clone CS0DE011VL04 of Placenta of Homo sapiens

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

(human).

CR621427
CR621427.1 GI:50502234
HTC; CNSLT cdNA.
Homo sapiens (human)

ORGANISM
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
REMARK

1 (bases 1 to 1592)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
Paradise Avenue
2 (bases 1 to 1592)

REFERENCE
AUTHORS
TITLE
JOURNAL

Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.

FEATURES
source

Location/Qualifiers
1..1592
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones="CS0DE011VL04"
/tissue_type="Placenta"
/plasmid="pCMVSPORT_6"

ORIGIN

Alignment Scores:

Pred. No.: 9,04e-167 Length: 1592
Score: 1460.50 Matches: 277
Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
DB: 3 Gaps: 1

US-10-698-228-1 (1-351) x CR621427 (1-1592)

Qy 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
Db 153 CCTGAGCGGACCCGCGCTCTGGCCAGCAAGACCGCGAGGAGGATCTTCCAGGAGCCAC 212
Qy 23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
Db 213 GGAGCCGAAACTAAAGCAGCTGCCCGCGGTGGAGGATGAGCGCTCTGAGAGAAA 272
Qy 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysG1 58
Db 273 CCCCCCGCGCTTTGTCTATCTTCCCATCGATGATATCTGGCAGATGATATAGAA 332
Qy 58 nAlaGlnAlaSerPheTyrPheTrpAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
Db 333 GGCAGAGCTTCTTTTGGACCGCGGAGGAGGTGGACCTCTCCAAGGACATTCAGCACATG 392
Qy 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAl 98
Db 393 GGAATCCCTGAAACCGCGAGGAGATATTTATATCCCATGTTCTTGGCTTCTTTCGAGC 452
Qy 98 aSerAspGlyIleValAsnGlnAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
Db 453 AAGCATGGCATAGTAATAAGAACTTGTGTGAGGATTTAGCCAAAGATTCAGATTAC 512
Qy 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuLeuGluAsnValHisSerGluMetTy 138
Db 513 AGAAGCCCGCTGTTTCTATGCTTCCAAATTTGCCATGGAACATACATCTTCTGAAATGTA 572


```

QY 258 aVallysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
935 TGTTGGATAGAACAGAGTTCTCTACTAGAGGCTTGCCCTGTGAAGCTCATTTGGATGAA 994
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuG1 298
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
995 TTGCACCTTAATGAAGCAATACATTGAGTTTGTGGCAGACAGACTTATGCTGGAACCTGGG 1054
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG1 318
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1055 TTTTACGAGTTTTCAGAGTAGAAGACCACTTTTGTACTTTATGGAGATATTTTCACTGGA 1114
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetal 338
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1115 AGGAAAGACTAACTTCTTTGAGAAGAGAGTAGGCGAGTATCAGAGGATGGGAGTGATGTC 1174
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1175 AAGTCCAAACAGAGAAATTTCTTTTACCTTGGATGCTGACTTC 1214
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 14
CR590959 1605 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CS0DE013YD22 of Placenta of Homo sapiens
DEFINITION
(human).
ACCESSION CR590959
VERSION CR590959.1 GI:50471766
KEYWORDS HTC; CDS; cDNA.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1605)
AUTHORS Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue
Genoscope.
2 (bases 1 to 1605)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source Location/Qualifiers
1..1605
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE013YD22"
/tissue_type="Placenta"
/plasmid="pCMVSPORT_6"
ORIGIN
Alignment Scores:
Pred. No.: 9.15e-167 Length: 1605
Score: 1460.50 Matches: 277
Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
DB: 3 Gaps: 1

US-10-698-228-1 (1-351) x CR590959 (1-1605)

QY 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

RESULT 15
CR614990
LOCUS

CR614990 1612 bp mRNA linear HTC 21-JUL-2004

```

150 CCTGAGCGGACCGCGCTCTCTGGCCAGCAAGACCGGAGGAGGATCTTCCAGGAGCCAC 209
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
210 GGAGCCGAAAACTAAAGCAGCTGCCCGCGGCTGGAGGATAGCCGCTGCTGAGAGAAA 269
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

38 rSerArgArgPheValIlePhePheProIleGlnTyrProAspIleTyrLysMetTyrLysG1 58
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
270 CCCCAGCGCTTGTCTATCTTCCCCTAGAGTACCATGATATCTGGCAGATGTATAAGAA 329
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

58 nAlaGlnAlaserPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
330 GGCAGAGCTTCTTTTGGACCGCGGAGAGGTGAGACTCTCCAGGACATTCAGACTG 389
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAl 98
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
390 GGAATCCTCGAAACCCGAGGAGAGATATTTTATATATATATATATATATATATATAT 449
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

98 aSerAspGlyIleValAsnGluLeuValGluArgPheSerGlnGluValGlnValPr 118
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
450 AAGCGATGGCATAGTAAATGAAAACTTGGTGGAGCGATTAGCCCAAGAAAGTTTCAATTAC 509
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
510 AGAAGCCCGCTGTCTTCTATGGCTTCCAAATTTGCCATGGAAACATACATCTCGAATGPA 569
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
570 TAGCTCTTATTATGACACTTACATAAAGATCCCAAGAAAGGGAATTTCTTCTCAATGC 629
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

158 aileGluThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTyrPheAlaAs 178
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
630 CATTGAAACGATGCTTGTGTCAAGAAAGAGGAGACTGGGCTTGCCTGGATTGGGGA 689
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

178 pArgLysSerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePh 198
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
690 CAAAGAGGCTACCTTGTGTAACGCTGTGTAGCCCTTGTCTCAGTGGAGGCAATTTCTT 749
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
750 TTCCGGTCTTTTTCGTCGATATTTCTGCTCAAGAAACGAGGACTGATGCTTGGCTCTAC 809
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
810 ATTTTCTAATGAATTTATTAGCAGAGATGAGGTTTACACTGTGATTTTGTCTGCTGAT 869
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleValAspAl 258
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
870 GTTCAACACCTGGTACACAAACCATCGGAGGAGAGAGTAAGAGAAAATAATATCAATGC 929
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
930 TGTTCCGATAGAACAGAGTTCTCTACTAGGCTTTCCTGCTGAGGCTCATTTGGATGAA 989
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluG1 298
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
990 TTGCACCTTAATGAAGCAATACATTGAGTTTGTGGCAGACAGACTTATGCTGGAACCTGG 1049
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG1 318
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1050 TTTTAGCAAGGTTTTCAGAGTAGAAGAACCCATTTTGTGATTTATGGAGATAATTTCACTGA 1109
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetal 338
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1110 AGGAAAGACTAACTTCTTTGAGAAGAGAGTAGGCGGAGTATCAGAGGATGGGAGTGATGTC 1169
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1170 AAGTCCAAACAGAGAAATTTCTTTTACCTTGGATGCTGACTTC 1209
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

DEFINITION full-length cDNA clone CS0DA009YB23 of Neuroblastoma of Homo sapiens (human).

ACCESSION CR614990

VERSION CR614990.1 GI:50495797

KEYWORDS HTC; CNSLT cDNA.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1612)

AUTHORS Li W.B., Gruber C., Jesse J. and Polayes D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished

REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600 Faraday Avenue

REFERENCE 2 (bases 1 to 1612)

AUTHORS Genoscope.

TITLE Direct Submission

JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : secref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.

FEATURES

source

1..1612

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DA009YB23"

/tissue type="Neuroblastoma"

/plasmid="pCMVSPORT_6"

ORIGIN

Alignment Scores:

Pred. No.: 9,21e-167 Length: 1612

Score: 1460.50 Matches: 277

Percent Similarity: 88.42% Conservative: 36

Best Local Similarity: 78.25% Mismatches: 35

Query Match: 80.20% Indels: 6

DB: 3 Gaps: 1

US-10-698-228-1 (1-351) x CR614990 (1-1612)

QY 4 ProGluArgProGluAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23

DB 152 CCTGAGCGGACCCCGCTCTGCGCCAGCAAGACCGGAGGAGGATCTTCCAGGAGCCAC 211

QY 23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38

DB 212 GGAGCGGAACCTAAAGCACTGCGCCGCGGAGGAGGATGAGCGCTGCTGAGAGAAA 271

QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTyrLysG1 58

DB 272 CCCCCCGCGCTTGTCTATCTTCCCATCGAGTACCAGATATCTGGCAGATGATAGAA 331

QY 58 nAlaGlnAlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78

DB 332 GGCAGAGGCTCTCTTTGGACCGCCGAGGAGGTGGACCTCTCCAAGGACATTCAGCACTG 391

QY 78 pAnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAl 98

DB 392 GGAATCCCTGAACCCGAGGAGAGATATTTATATCCCATGTTCTGCGCTTCTTTGCGAG 451

QY 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118

DB 452 AAGCGATGGCATAGTAATGAACACTTGGTGGAGCGATTAGCCAGAGATTCAGATTAC 511

QY 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138

512 AGAAGCCCGCTGTTTCTATGGCTTCCAAATGGCCATGGAAACATACATTCTGAAATGTA 571

138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158

572 TAGTCTTCTTATTGACACTTACATAAAAGATCCCAAGAAAGGGAATTTCTCTCAATGC 631

158 aIleGluThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178

632 CATTGAAACCATGCTTGTGTCAAGAGAAGGAGACTGGGCCCTTGGCTGTGATTTGGGA 691

178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePh 198

692 CAAGAGGCTACCTATGCTGNAACGTTGTAGCTTTGCTCAGTGGAGGATTTCTTCT 751

198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218

752 TTCGGTTCCTTTTGGCTGCTATTTCTGCTCAAGAAACGAGACTGATGCTGGCTCAC 811

218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238

812 ATTTTCTTAATCAACTTATTAGCAGAGATGAGGGTTTACACTGTGATTTTGTCTGCTGAT 871

238 tPheGluThrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258

872 GTTCAACACCTGCTGACAAACCATCGAGGAGAGAGTAGAGAAATAATTTATCAATGC 931

258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278

932 TGTTCCGATAGAACAGGAGTTCTCTCACTGAGGCCCTTGGCTGTGAAGCTCATTTGGGATGAA 991

278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAlaAspArgLeuLeuValGluLeuCl 298

992 TTGCACCTTAATGAAGCAATACATTTGTTGTGCGACAGACTTATGCTGGAACCTGGG 1051

298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGl 318

1052 TTTTAGCAAGTTTTCAGAGTAGAGAACCCATTTGACTTTATGGAGAATATTTTCACTGGA 1111

318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338

1112 AGGAAAGACTAACTTCTTTGAGAAGAGAGTAGGCGAGTATCAGAGGATGGGAGTGATGTC 1171

338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351

1172 AAGTCCAACACAGAAATTTCTTTTACCTTGGATGCTGACTTC 1211

Search completed: June 16, 2005, 14:11:19
Job time : 3873 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2005, 06:52:16 ; Search time 4888 Seconds
(without alignments)
10438.491 Million cell updates/sec

Title: US-10-698-228-2
Perfect score: 1053
Sequence: 1 atggggcaccgaaagcc.....tcaccttgatgcagatttt 1053

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1053	100.0	1053	6	AR454868
2	1053	100.0	1053	6	BD064764
3	1053	100.0	1053	6	BD093077
4	1053	100.0	1056	6	CQ714252
5	1053	100.0	1056	9	AB166671
6	1053	100.0	4955	6	AR454869
7	1053	100.0	4955	6	BD064765
8	1053	100.0	4955	6	BD093078
9	1053	100.0	4955	9	AB036063
10	1051.4	99.8	1053	6	AR454877
11	1051.4	99.8	1053	6	BD064773
12	1051.4	99.8	1053	6	BD093086
13	1051.4	99.8	1081	6	AR454870
14	1051.4	99.8	1081	6	BD064766
15	1051.4	99.8	1081	6	BD093079
16	1049.8	99.7	1601	6	BD156916
17	1049.8	99.7	1601	6	AX877905
18	1049.8	99.7	1601	9	AK001965
19	895.2	85.0	4532	10	BC058103

20	855	81.2	4571	9	BC042468	BC042468 Homo sapi
21	853.4	81.0	900	9	AB163437	AB163437 Homo sapi
22	837.6	79.5	171737	10	AC122379	AC122379 Mus muscu
23	579	55.0	2500	6	AR300103	AR300103 Sequence
24	579	55.0	2500	6	AR438503	AR438503 Sequence
25	579	55.0	2500	6	AR562011	AR562011 Sequence
26	578.4	54.9	1649	9	BC001886	BC001886 Homo sapi
27	578.4	54.9	1653	9	BC030154	BC030154 Homo sapi
28	578.4	54.9	1794	9	S40301	S40301 Homo sapien
29	578.4	54.9	2500	6	AR380825	AR380825 Sequence
30	578.4	54.9	2500	6	AX333242	AX333242 Sequence
31	578.4	54.9	2500	6	AX333687	AX333687 Sequence
32	578.4	54.9	2500	6	AX334345	AX334345 Sequence
33	578.4	54.9	2500	9	HSRR2SS	XS9618 H.sapiens R
34	578.4	54.9	3039	9	AK092671	AK092671 Homo sapi
35	578.2	54.9	1562	5	BC061353	BC061353 Xenopus t
36	574.6	54.6	1538	5	CR762018	CR762018 Xenopus t
37	574.6	54.6	1611	5	CR848103	CR848103 Xenopus t
38	574.6	54.6	1647	5	BC080161	BC080161 Xenopus t
39	573.6	54.5	217774	2	AC120603	AC120603 Rattus no
40	570.4	54.2	3203	9	AK123010	AK123010 Homo sapi
41	570.4	54.2	2111	10	MUSNRM2A	M14223 Mouse ribon
42	568.8	54.0	1170	6	E14825	E14825 cDNA encodi
43	566.2	53.8	1570	5	BC072071	BC072071 Xenopus l
44	562.6	53.4	2057	5	BC044355	BC044355 Danio rer
45	561.8	53.4	1406	5	CR353642	CR353642 Gallus ga

ALIGNMENTS

RESULT 1
LOCUS AR454868 1053 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 2 from patent US 6682917.
ACCESSION AR454868
VERSION AR454868.1 GI:42688823
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1053)
AUTHORS Nakamura,Y., Arakawa,H. and Tanaka,H.
TITLE Protein having a ribonucleotide Reductase activity and a DNA thereof
JOURNAL Patent: US 6682917-A 2 27-JAN-2004;
FEATURES Location/Qualifiers
source 1..1053
/organism="unknown"
/mol_type="genomic DNA"

Qy	1	ATGGGGCACC	CGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA	60
Db	1	ATGGGGCACC	CGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA	60
Qy	61	GACACCAAC	GAAAGTCAATGAAGAGCCACTCCTTAAGAAAGAGTTCTCGC	120
Db	61	GACACCAAC	GAAAGTCAATGAAGAGCCACTCCTTAAGAAAGAGTTCTCGC	120
Qy	121	CGGTTTGAT	CTTTTCCAATCCAGTACCTCGATATTTGGAAATATATAAAGGACAG	180
Db	121	CGGTTTGAT	CTTTTCCAATCCAGTACCTCGATATTTGGAAATATATAAAGGACAG	180
Qy	181	GCCTTCCT	TCGGACAGCAGAGAGGTCCACTTATCAAAAGGATCTCCCTCAGTGAACAG	240
Db	181	GCCTTCCT	TCGGACAGCAGAGAGGTCCACTTATCAAAAGGATCTCCCTCAGTGAACAG	240
Qy	241	CTTAAGAG	CAGATGAGAAGTACTTCTCATCTTACATCTTGTAGCCTTTTGTGACCCAGTGTAT	300

Db 241 CTTAAAGCAGATGAGAAGTACTTCATCTCTCATCTTACGCTTTTTCAGCAGGTAT 300
 QY 301 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
 Db 301 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
 QY 361 CGCTGTTTCTATGGCTTTCAAAATTCATCAGAGAAATTTTCATCAGAGATGTACAGTTG 420
 Db 361 CGCTGTTTCTATGGCTTTCAAAATTCATCAGAGAAATTTTCATCAGAGATGTACAGTTG 420
 QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTATTAATGCAATGAA 480
 Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTATTAATGCAATGAA 480
 QY 481 ACCATGCCCTATCTTAAGAAAAGCAGATTTGGCCCTTGGATGGATAGCAGATAGAAA 540
 Db 481 ACCATGCCCTATCTTAAGAAAAGCAGATTTGGCCCTTGGATGGATAGCAGATAGAAA 540
 QY 541 TCTACTTTTGGGAAAAGAGTGGTGGCTTTGCTGTGTAAGAGAGTGTTCCTCAGGA 600
 Db 541 TCTACTTTTGGGAAAAGAGTGGTGGCTTTGCTGTGTAAGAGAGTGTTCCTCAGGA 600
 QY 601 TCTTTTGTCTATATCTGGCTTAAGAAAGAGGCTTTATGCGAGACTCACTTTTTC 660
 Db 601 TCTTTTGTCTATATCTGGCTTAAGAAAGAGGCTTTATGCGAGACTCACTTTTTC 660
 QY 661 AATGAACCTCATCAGCAGATGAAGACTTCATCTGACCTTTGCTTGCCTGATGTTCCAA 720
 Db 661 AATGAACCTCATCAGCAGATGAAGACTTCATCTGACCTTTGCTTGCCTGATGTTCCAA 720
 QY 721 TACTTAGTAATAAGCTTCAGAAAGAGGCTCAGGAGATCATTTGTGATGCTGTCAAA 780
 Db 721 TACTTAGTAATAAGCTTCAGAAAGAGGCTCAGGAGATCATTTGTGATGCTGTCAAA 780
 QY 781 ATTGACGAGAGTTTTTAACAGAGCTTCCAGTGGCTCATTTGGAATGAATTCATT 840
 Db 781 ATTGACGAGAGTTTTTAACAGAGCTTCCAGTGGCTCATTTGGAATGAATTCATT 840
 QY 841 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTAATCTTGGAATTCATGATCTCA 900
 Db 841 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTAATCTTGGAATTCATGATCTCA 900
 QY 901 AAGTTTTTTCAGCAGAAAATCCTTTTGAATTTATGAAAACATTTCTTTAGAGGAAA 960
 Db 901 AAGTTTTTTCAGCAGAAAATCCTTTTGAATTTATGAAAACATTTCTTTAGAGGAAA 960
 QY 961 ACAATTTCTTTGAGAAAAGTTCAGAGTATCAGGTTTTGCGATTTATGCGAGAAACC 1020
 Db 961 ACAATTTCTTTGAGAAAAGTTCAGAGTATCAGGTTTTGCGATTTATGCGAGAAACC 1020
 QY 1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053
 Db 1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053

RESULT 2
 BD064764
 LOCUS
 DEFINITION Novel protein and DNA thereof.
 ACCESSION BD064764
 VERSION BD064764.1 GI:22610367
 KEYWORDS JP 2001269184-A/1.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1053)
 Nakamura, Y., Arakawa, H. and Tanaka, H.
 Novel protein and DNA thereof
 Patent: JP 2001269184-A 1 02-OCT-2001;
 YUSUKE NAKAMURA, TAKEDA CHEMICAL INDUSTRIES LTD
 OS Homo sapiens (human)

PN JP 2001269184-A/1
 PD 02-OCT-2001
 PF 27-JUN-2000 JP 2000192401
 PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
 C12N15/09, A61K31/711, A61K38/43, A61K45/00, A61K48/00, PC
 A61P35/00,
 PC A61P43/00, C07K16/32, C07K16/40, C12N1/15, C12N1/19, C12N1/21 PC
 , C12N5/10, C12N9/02,
 PC C12P21/02, C12P21/08, G01N33/15, G01N33/50, G01N33/53// (C12N15/09,
 C12R1:91),
 PC (C12N15/09, C12R1:19), (C12N9/02, C12R1:91), (C12P21/02, C12R1:91),
 PC (C12P21/08, C12R1:91), C12N15/00, A61K37/02, A61K37/48, C12N5/00,
 PC (C12N15/00, C12R1:91), C12N15/00, C12R1:19)
 CC Novel protein and DNA thereof
 FH Key Location/Qualifiers
 FT source 1..1053
 /organism='Homo sapiens (human)'.
 FEATURES
 source
 Location/Qualifiers
 1..1053
 /organism='Homo sapiens'
 /mol_type='genomic DNA'
 /db_xref='taxon:9606'
 ORIGIN
 Query Match 100.0%; Score 1053; DB 6; Length 1053;
 Best Local Similarity 100.0%; Pred. No. 1.2e-255;
 Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGGCGCACCGGAAAGCGCGAGCGCGGGCTGGATCAGGATGAGAGATCATCTTCA 60
 Db 1 ATGGCGCACCGGAAAGCGCGAGCGCGGGCTGGATCAGGATGAGAGATCATCTTCA 60
 QY 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
 Db 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
 QY 121 CGGTTTGTATCTTTCCAAATCCAGTACCTGTATATTTGGAAAATGTAATAACAGGACAG 180
 Db 121 CGGTTTGTATCTTTCCAAATCCAGTACCTGTATATTTGGAAAATGTAATAACAGGACAG 180
 QY 181 GCTTCCTTCTGGACAGCAGAGAGGTGAGTATCAAGAGGATCTCCCTCACTGGAAACAG 240
 Db 181 GCTTCCTTCTGGACAGCAGAGAGGTGAGTATCAAGAGGATCTCCCTCACTGGAAACAG 240
 QY 241 CTTAAAGCAGATGAGAAGTACTTCTCTCACATCTTAGCCTTTTTCAGGCCAGTGTAT 300
 Db 241 CTTAAAGCAGATGAGAAGTACTTCTCTCACATCTTAGCCTTTTTCAGGCCAGTGTAT 300
 QY 301 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
 Db 301 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
 QY 361 CGCTGTTTCTATGGCTTTCAAAATTCATCAGAGAAATTTTCATCAGAGATGTACAGTTG 420
 Db 361 CGCTGTTTCTATGGCTTTCAAAATTCATCAGAGAAATTTTCATCAGAGATGTACAGTTG 420
 QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTATTAATGCAATGAA 480
 Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTATTAATGCAATGAA 480
 QY 481 ACCATGCCCTATGTTAAGAAAAGAGAGATTTGGCCCTTGGATGGATAGCAGATAGAAA 540
 Db 481 ACCATGCCCTATGTTAAGAAAAGAGAGATTTGGCCCTTGGATGGATAGCAGATAGAAA 540
 QY 541 TCTACTTTTGGGAAAAGAGTGGTGGCTTTGCTGTGTAAGAGAGTGTTCCTCAGGA 600
 Db 541 TCTACTTTTGGGAAAAGAGTGGTGGCTTTGCTGTGTAAGAGAGTGTTCCTCAGGA 600
 QY 601 TCTTTTGTCTATATCTGGCTTAAGAAAAGAGGCTTTATGCGAGACTCACTTTTTC 660
 Db 601 TCTTTTGTCTATATCTGGCTTAAGAAAAGAGGCTTTATGCGAGACTCACTTTTTC 660
 QY 661 AATGAACCTCATCAGCAGATGAAGACTTCATCTGACCTTTGCTTGCCTGATGTTCCAA 720

[illegible]

```
VERSION CQ714252.1 GI:42275109
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 186 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES Location/Qualifiers
source 1..1056
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 1053; DB 6; Length 1056;
Best Local Similarity 100.0%; Pred. No. 1.2e-255;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCGACCCCGAAGAGCGGCGGCTGGATCAGATGAGAGATCATCTTCA 60
DB 1 ATGGCGACCCCGAAGAGCGGCGGCTGGATCAGATGAGAGATCATCTTCA 60
QY 61 GACACCAAGCAAGTCAATTAAGTCAATGATGAGAGCCACTCTTAAGAAAGAGTTCGCG 120
DB 61 GACACCAAGCAAGTCAATTAAGTCAATGATGAGAGCCACTCTTAAGAAAGAGTTCGCG 120
QY 121 CGTTTGTTCATCTTCCATCCAGTACCCTGATATTTGGAAAAATGTATAAACAGGCACAG 180
DB 121 CGTTTGTTCATCTTCCATCCAGTACCCTGATATTTGGAAAAATGTATAAACAGGCACAG 180
QY 181 GCTTCCTCTGGACAGCAGAAAGGTCGACTTATCAAAAGATCTCCCTCACTGGAACAAG 240
DB 181 GCTTCCTCTGGACAGCAGAAAGGTCGACTTATCAAAAGATCTCCCTCACTGGAACAAG 240
QY 241 CTTAAAGCAGATGAGAAGTACTTCATCTCTCACATCTTAGCCCTTTTTCGAGCCAGTGAT 300
DB 241 CTTAAAGCAGATGAGAAGTACTTCATCTCTCACATCTTAGCCCTTTTTCGAGCCAGTGAT 300
QY 301 GGAATGTAAATGAAAATTTGGTGGAGCGCTTTTAGTCAGAGGTGCGAGGGCT 360
DB 301 GGAATGTAAATGAAAATTTGGTGGAGCGCTTTTAGTCAGAGGTGCGAGGGCT 360
QY 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAATGTTTCACCTCAGAGATGTACAGTTTG 420
DB 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAATGTTTCACCTCAGAGATGTACAGTTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATTGAA 480
DB 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATTGAA 480
QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATGCGCCTTCGATGATAGCAGATAGAA 540
DB 481 ACCATGCCCTATGTTAAGAAAAGCAGATGCGCCTTCGATGATAGCAGATAGAA 540
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAAGAGTTCCTCTCAGGA 600
DB 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAAGAGTTCCTCTCAGGA 600
QY 601 TCTTTTGTGCTATATCTCGCTAAAGAGAGAGGTCTTATGCAAGACTCATTCTTTC 660
DB 601 TCTTTTGTGCTATATCTCGCTAAAGAGAGAGGTCTTATGCAAGACTCATTCTTTC 660
QY 661 AATGAATCATCAGCAGAGATGAGGACTTCATCTGTCATCTTCTGCTGCTGATGTTCCAA 720
DB 661 AATGAATCATCAGCAGAGATGAGGACTTCATCTGTCATCTTCTGCTGCTGATGTTCCAA 720
QY 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 780
|||||
721 TACTTAGTAAATAAGCCTTCAGAAAGAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 780
|||||
781 ATTGACGACGAGTTCCTTTTAAACAGAAAGCCTTCGCCAGTTGGCCTCATTTGAAATGAATTGCATT 840
781 ATTGACGACGAGTTCCTTTTAAACAGAAAGCCTTCGCCAGTTGGCCTCATTTGAAATGAATTGCATT 840
841 TTGATGAAAACAGTACATTCGAGTTTGTAGCTGACAGATTAATCTTGTGGAACCTTGGATTCTCA 900
841 TTGATGAAAACAGTACATTCGAGTTTGTAGCTGACAGATTAATCTTGTGGAACCTTGGATTCTCA 900
901 AAGTTTTTTCAGGCAGAAAATCCTTTTGTATTTATGAAAACATTTCTTTAGAAAGGAAA 960
901 AAGTTTTTTCAGGCAGAAAATCCTTTTGTATTTATGAAAACATTTCTTTAGAAAGGAAA 960
961 ACAAAATTTCTTTGAGAAAACGAGTTTCAGAGTATCAGCGTTTTCAGTGTATGGCAGAAACC 1020
961 ACAAAATTTCTTTGAGAAAACGAGTTTCAGAGTATCAGCGTTTTCAGTGTATGGCAGAAACC 1020
1021 ACAGATAACGCTTTCACCTTTGGATGCAGATTTT 1053
1021 ACAGATAACGCTTTCACCTTTGGATGCAGATTTT 1053
RESULT 5
ABI166671 1056 bp mRNA linear PRI 09-MAR-2004
LOCUS Homo sapiens p53R2 mRNA for p53-inducible ribonucleotide reductase
DEFINITION small subunit 2, complete cds.
ACCESSION ABI166671
VERSION ABI166671.1 GI:45259568
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Ugai,H. and Yokoyama,K.K.
TITLE Homo sapiens p53R2 mRNA for p53-inducible ribonucleotide reductase
JOURNAL Published Only in Database (2004)
REFERENCE 2 (bases 1 to 1056)
AUTHORS Ugai,H. and Yokoyama,K.K.
TITLE Direct Submission
JOURNAL Submitted (05-MAR-2004) Hideyo Ugai, RIKEN Bioresource Center, Gene
Engineering Division; 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074,
Japan (E-mail:ugai@rtc.riken.go.jp, Tel:81-29-836-3612,
Fax:81-29-836-9120)
FEATURES Location/Qualifiers
source 1..1056
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="p53R2.C19"
/cell_line="Saos-2"
/cell_type="Osteogenic sarcoma"
1..1056
/gene="p53R2"
/gene="p53R2"
/codon_start=1
/product="p53-inducible ribonucleotide reductase small
subunit 2"
/protein_id="BAD12267.1"
/db_xref="GI:45259568"
/translation="MGDPPEAAGLQDERSSDTNESEIKSNEEPLLRKSRRFVI
FPYQPDIDWMYKQAQAFWBEVDLSKDLPHWNKLKADKEYFISHLAFPAASDGI
VNEVLRFQEQVQPEARCFYGFQIILNVHSEMSLLIDTYIRDPKRFLFNAIE
TWPYKKADMALRWIADRKSTGERVVAFAAVEGVFSGSFAAI FWLKRGLMPLT
FSNELLSRDEGLHCDFACLMFOILVINKPSEERVELIIVADVAKIEQELTALPYGLIG
MNCILMKQYIEFVAIDRLVLVELGFSKVFQENPFDMENI SLEGTINFTEKRVSEYQRF
AVMAETTDNVFTLDAF"
ORIGIN
```

Query Match 100.0%; Score 1053; DB 9; Length 1056;
 Best Local Similarity 100.0%; Pred. No. 1.2e-255;
 Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGCGACCGGAAAGCGCGGAGCGCGGCTGGATCAGATGAGATCATCTTCA 60
 DB 1 ATGGGCGACCGGAAAGCGCGGAGCGCGGCTGGATCAGATGAGATCATCTTCA 60

QY 61 GACACCAACGAAAGTAAAGTCAAAATGAAGAGCCACTCCCTAAGAAAGTCTCGC 120
 DB 61 GACACCAACGAAAGTAAAGTCAAAATGAAGAGCCACTCCCTAAGAAAGTCTCGC 120

QY 121 CGGTTTGTCACTTTCAATCCAGTACCTGATATTGGAAATGATATAACAGGCACAG 180
 DB 121 CGGTTTGTCACTTTCAATCCAGTACCTGATATTGGAAATGATATAACAGGCACAG 180

QY 181 GCTTCCTCTGGACAGACAGAGGTCGACTTATCAAGAGATCTCCCTCACTGGAACAAG 240
 DB 181 GCTTCCTCTGGACAGACAGAGGTCGACTTATCAAGAGATCTCCCTCACTGGAACAAG 240

QY 241 CTTAAAGCAGATGAGAAAGTACTTCACTCTCACATCTTACGCTTTTGGCAGCCAGTAT 300
 DB 241 CTTAAAGCAGATGAGAAAGTACTTCACTCTCACATCTTACGCTTTTGGCAGCCAGTAT 300

QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
 DB 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360

QY 361 CGGTGTTTCTATGGCTTTCAAAATCTCATCGAGAAATGTTCACTCAGAGATGTAAGTTTG 420
 DB 361 CGGTGTTTCTATGGCTTTCAAAATCTCATCGAGAAATGTTCACTCAGAGATGTAAGTTTG 420

QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 480
 DB 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 480

QY 481 ACCATGCCCTATGTTAAAGAAAGACAGATTTGGCGCTTGGCATGGATAGAGATAGAAA 540
 DB 481 ACCATGCCCTATGTTAAAGAAAGACAGATTTGGCGCTTGGCATGGATAGAGATAGAAA 540

QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGCTAGAGAGGTTTCTCTCAGGA 600
 DB 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGCTAGAGAGGTTTCTCTCAGGA 600

QY 601 TCTTTTGTGCTATATTCTGGCTAAAGAGAGGTCCTTATGCCAGGACTCACTTTTTC 660
 DB 601 TCTTTTGTGCTATATTCTGGCTAAAGAGAGGTCCTTATGCCAGGACTCACTTTTTC 660

QY 661 AATGAACCTCATCAGCAGATGAAGGACTTCTACGTGACCTTTGCTTGGCTGATGTTCCAA 720
 DB 661 AATGAACCTCATCAGCAGATGAAGGACTTCTACGTGACCTTTGCTTGGCTGATGTTCCAA 720

QY 721 TACTTAGTAAATAAGCTTTCAAGAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 780
 DB 721 TACTTAGTAAATAAGCTTTCAAGAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 780

QY 781 ATTGAGCAGAGGTTTTTAACAGAGCCTTGGCAGTTGGCCCTCATTTGGAATGAATTCGATT 840
 DB 781 ATTGAGCAGAGGTTTTTAACAGAGCCTTGGCAGTTGGCCCTCATTTGGAATGAATTCGATT 840

QY 841 TTGATCAAAACAGTACATTGAGTTTGTAGCTGACAGATTAAGTGGAACTTGGATTCTCA 900
 DB 841 TTGATCAAAACAGTACATTGAGTTTGTAGCTGACAGATTAAGTGGAACTTGGATTCTCA 900

QY 901 AAGGTTTTTTCAGGAGAAATCTTTTGTATTTTATGGAATACTTTTATAGAGGAA 960
 DB 901 AAGGTTTTTTCAGGAGAAATCTTTTGTATTTTATGGAATACTTTTATAGAGGAA 960

QY 961 ACAAAATTTCTTTGAGAAAACAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020
 DB 961 ACAAAATTTCTTTGAGAAAACAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020

QY 1021 ACAGATAACGCTCTTCACTTGGATGAGATTTT 1053
 DB 1021 ACAGATAACGCTCTTCACTTGGATGAGATTTT 1053

RESULT 6
 AR454869
 LOCUS AR454869 4955 bp DNA linear PAT 20-FEB-2004
 DEFINITION Sequence 3 from patent US 6682917.
 ACCESSION AR454869
 VERSION AR454869.1 GI:42688824
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 4955)
 AUTHORS Nakamura, Y., Arakawa, H. and Tanaka, H.
 TITLE Protein having a ribonucleotide Reductase activity and a DNA thereof
 JOURNAL Patent: US 6682917-A 3 27-JAN-2004;
 FEATURES Location/Qualifiers
 source 1..4955
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN
 Query Match 100.0%; Score 1053; DB 6; Length 4955;
 Best Local Similarity 100.0%; Pred. No. 1.1e-255;
 Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGCGACCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGATCATCTTCA 60
 DB 245 ATGGGCGACCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGATCATCTTCA 304

QY 61 GACACCAACGAAAGTAAAGTCAAAATGAAGAGCCACTCCCTAAGAAAGTCTCGC 120
 DB 305 GACACCAACGAAAGTAAAGTCAAAATGAAGAGCCACTCCCTAAGAAAGTCTCGC 364

QY 121 CGGTTTGTGCTATTTCCAAATCCAGTACCTGATATTGGAAATGATATAACAGGCACAG 180
 DB 365 CGGTTTGTGCTATTTCCAAATCCAGTACCTGATATTGGAAATGATATAACAGGCACAG 424

QY 181 GCTTCTCTTGGACAGCAGAGAGGTGCGACTTATCAAGAGATCTCCCTCACTGGAACAAG 240
 DB 425 GCTTCTCTTGGACAGCAGAGAGGTGCGACTTATCAAGAGATCTCCCTCACTGGAACAAG 484

QY 241 CTTAAAGCAGATGAGAAAGTACTTCACTCTCACATCTTACGCTTTTGGCAGCCAGTAT 300
 DB 485 CTTAAAGCAGATGAGAAAGTACTTCACTCTCACATCTTACGCTTTTGGCAGCCAGTAT 544

QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
 DB 545 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 604

QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGATGTTCACTCAGAGATGTAAGTTTG 420
 DB 605 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGATGTTCACTCAGAGATGTAAGTTTG 664

QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 480
 DB 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 724

QY 481 ACCATGCCCTATGTTAAAGAAAGACAGATTTGGCGCTTGGCATGGATAGAGATAGAAA 540
 DB 725 ACCATGCCCTATGTTAAAGAAAGACAGATTTGGCGCTTGGCATGGATAGAGATAGAAA 784

QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGCTAGAGAGGTTTCTTCTCAGGA 600
 DB 785 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGCTAGAGAGGTTTCTTCTCAGGA 844

QY 601 TCTTTTGTGCTATATTCTGGCTAAAGAGAGGTCCTTATGCCAGGACTCACTTTTTC 660
 DB 845 TCTTTTGTGCTATATTCTGGCTAAAGAGAGGTCCTTATGCCAGGACTCACTTTTTC 904

```

QY 661 AATGAACATCATCAGCAGAGATGAAGACATTCACCTGTGACTTTGCTTGCCTGATGTTCCAA 720
DB 905 AATGAACATCATCAGCAGAGATGAAGACATTCACCTGTGACTTTGCTTGCCTGATGTTCCAA 964
QY 721 TACTTAGTAAATAAGCCTTCAGAGAAAGGTCAGGAGATCAATGTTGATGCTGTCAAA 780
DB 965 TACTTAGTAAATAAGCCTTCAGAGAAAGGTCAGGAGATCAATGTTGATGCTGTCAAA 1024
QY 781 ATTGACGAGAGTATTTTAAACAGAGCCTTCCAGTGTGCTCATTTGGAATGAATTCATTT 840
DB 1025 ATTGACGAGAGTATTTTAAACAGAGCCTTCCAGTGTGCTCATTTGGAATGAATTCATTT 1084
QY 841 TTGATGAACAGTACATGATGTTGATGTTGAGTGAAGATTTCTGTGGAACCTTGGATTTCA 900
DB 1085 TTGATGAACAGTACATGATGTTGATGTTGAGTGAAGATTTCTGTGGAACCTTGGATTTCA 1144
QY 901 AAGGTTTTTTCAGCAGAGAAATCCTTTTGAATTTTATGGAACATTTCTTTAGAGGAAA 960
DB 1145 AAGGTTTTTTCAGCAGAGAAATCCTTTTGAATTTTATGGAACATTTCTTTAGAGGAAA 1204
QY 961 ACAAATTTCTTTCAGAAACAGATTTTCAGAGTATCAGCGTTTTCGAGTTATGGCAGAAACC 1020
DB 1205 ACAAATTTCTTTCAGAAACAGATTTTCAGAGTATCAGCGTTTTCGAGTTATGGCAGAAACC 1264
QY 1021 ACAGATAACGTTCTTCACTTGGATGCGAGATTTT 1053
DB 1265 ACAGATAACGTTCTTCACTTGGATGCGAGATTTT 1297

```

RESULT 7

```

BD064765
LOCUS Novel protein and DNA thereof. 4955 bp DNA linear PAT 27-AUG-2002
DEFINITION Novel protein and DNA thereof.
ACCESSION BD064765
VERSION BD064765.1 GI:22610368
KEYWORDS JP 2001269184-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 4955)
Nakamura,Y., Arakawa,H. and Tanaka,H.
Novel protein and DNA thereof
Patent: JP 2001269184-A 2 02-OCT-2001;
YUSUKE NAKAMURA, TAKEDA CHEMICAL INDUSTRIES LTD
OS Homo sapiens (human)
PN JP 2001269184-A/2
PD 02-OCT-2001
PF 27-JUN-2000 JP 2000192401
PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
C12N15/09,A61K31/711,A61K38/00,A61K38/43,A61K45/00,A61K48/00, PC
A61P35/00,
PC A61P43/00,C07K16/32,C07K16/40,C12N1/15,C12N1/19,C12N1/21 PC
C12N5/10,C12N9/02,
PC C12P21/02,C12P21/08,G01N33/15,G01N33/50,G01N33/53///C12N15/09,
PC C12R1.91,
PC C12N15/09,C12R1.19),(C12N9/02,C12R1.91),(C12P21/02,C12R1.91),
PC (C12P21/08,C12R1.91),C12N15/00,A61K37/02,A61K37/48,C12N5/00,
PC (C12N15/00,C12R1.91),(C12N15/00,C12R1.19)
CC Novel protein and DNA thereof
FH Key Location/Qualifiers
FT source 1.4955
FT Location/Qualifiers
1.4955
/organism="Homo sapiens (human)".

```

FEATURES

```

source
1.4955
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 1053; DB 6; Length 4955;
Best Local Similarity 100.0%; Pred. No. 1.1e-255;

```

```

Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCGACCGCGAAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
DB 245 ATGGCGACCGCGAAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
QY 61 GACACCAACGAAGTGAATTAAGTCAANTGAAGGCCACTTCTTAAGAAAGATTTCTCGC 120
DB 305 GACACCAACGAAGTGAATTAAGTCAANTGAAGGCCACTTCTTAAGAAAGATTTCTCGC 364
QY 121 CGGTTTCTCATCTTTCATATCCAGTACCTCATATTTGGAAAAATGTATAAACAGGCACAG 180
DB 365 CGGTTTCTCATCTTTCATATCCAGTACCTCATATTTGGAAAAATGTATAAACAGGCACAG 424
QY 181 GCTTCTCTTGGACAGCAGAGAGGTCACTTATCAAAAGGATCTCCCTCAGTGGAAACAG 240
DB 425 GCTTCTCTTGGACAGCAGAGAGGTCACTTATCAAAAGGATCTCCCTCAGTGGAAACAG 484
QY 241 CTTAAAGCAGATGAAGTACTTTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTGTAT 300
DB 485 CTTAAAGCAGATGAAGTACTTTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTGTAT 544
QY 301 GGAATCTGTAATGAATTTTGGTGGAGCGCTTTAGTCAGGAGGTGACGTTCCAGAGGCT 360
DB 545 GGAATCTGTAATGAATTTTGGTGGAGCGCTTTAGTCAGGAGGTGACGTTCCAGAGGCT 604
QY 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCAGAGATGTTTCACTCAGAGATGTACAGTTTG 420
DB 605 CGCTGTTTCTATGGCTTTCAAATTTCTCATCAGAGATGTTTCACTCAGAGATGTACAGTTTG 664
QY 421 CTGATAGACATTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAATGCAATTGAA 480
DB 665 CTGATAGACATTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAATGCAATTGAA 724
QY 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGGCTTTGCGATGATGATAGCAGATAGAAA 540
DB 725 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGGCTTTGCGATGATGATAGCAGATAGAAA 784
QY 541 TCTACTTTTGGGAAAAAGAGTGGTGGCTTTGCTGTGTAGAAAGGAGTTTCTTCTCAGGA 600
DB 785 TCTACTTTTGGGAAAAAGAGTGGTGGCTTTGCTGTGTAGAAAGGAGTTTCTTCTCAGGA 844
QY 601 TCTTTTGTGCTATATTTCTGTGCTAAAGAGAGGTCTTATGCGCAGGACTCTCATCTTTTCC 660
DB 845 TCTTTTGTGCTATATTTCTGTGCTAAAGAGAGGTCTTATGCGCAGGACTCTCATCTTTTCC 904
QY 661 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
DB 905 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 964
QY 721 TACTTAGTAAATAAGCCTTCAGAGAAAGGTCAGGAGATCAATGTTGATGCTGTCAAA 780
DB 965 TACTTAGTAAATAAGCCTTCAGAGAAAGGTCAGGAGATCAATGTTGATGCTGTCAAA 1024
QY 781 ATTGACGAGAGTATTTTAAACAGAGCCTTCCAGTGTGCTCATTTGGAATGAATTCATTT 840
DB 1025 ATTGACGAGAGTATTTTAAACAGAGCCTTCCAGTGTGCTCATTTGGAATGAATTCATTT 1084
QY 841 TTGATGAACAGTACATGATGTTTGTAGTGAAGATTTCTGTGGAACCTTGGATTTCA 900
DB 1085 TTGATGAACAGTACATGATGTTTGTAGTGAAGATTTCTGTGGAACCTTGGATTTCA 1144
QY 901 AAGGTTTTTTCAGCAGAGAAATCCTTTTGAATTTTATGGAACATTTCTTTAGAGGAAA 960
DB 1145 AAGGTTTTTTCAGCAGAGAAATCCTTTTGAATTTTATGGAACATTTCTTTAGAGGAAA 1204
QY 961 ACAAATTTCTTTCAGAAACAGATTTTCAGAGTATCAGCGTTTTCGAGTTATGGCAGAAACC 1020
DB 1205 ACAAATTTCTTTCAGAAACAGATTTTCAGAGTATCAGCGTTTTCGAGTTATGGCAGAAACC 1264
QY 1021 ACAGATAACGTTCTTCACTTGGATGCGAGATTTT 1053
DB 1265 ACAGATAACGTTCTTCACTTGGATGCGAGATTTT 1297

```

```

RESULT 8
BD093078
LOCUS
DEFINITION Novel protein and its DNA.
ACCESSION BD093078
VERSION BD093078.1 GI:22638666
KEYWORDS WO 0100799-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 4955)
Nakamura, Y., Arakawa, H. and Tanaka, H.
Novel protein and its DNA
Patent: WO 0100799-A 2 04-JAN-2001;
TAKEDA CHEMICAL INDUSTRIES LTD, YUSUKE NAKAMURA, HIROFUMI ARAKAWA,
HIROSHI TANAKA
OS Homo sapiens (human)
PN WO 0100799-A/2
PD 04-JAN-2001
PF 27-JUN-2000 WO 2000JP004189
PR 28-JUN-1999 JP 99P 181131, 06-JUL-1999 JP 99P 192391 PR
21-JAN-2000 JP 00P 017770
PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
C12N9/04, C12N15/53, C12N1/15, C12N1/19, C12N1/21, C12N5/10 PC
.C12P21/02, A61K38/44,
PC A61K48/00, C07K16/40, G01N33/53, C12Q1/26
CC
FH Key Location/Qualifiers.
FEATURES
source
1..4955
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 1053; DB 6; Length 4955;
Best Local Similarity 100.0%; Pred. No. 1.1e-255;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGCGACCGGAAAGCGCGGCGGCGGTGGATCAGGATGAGAGATCATCTTCA 60
DB 245 ATGGGCGACCGGAAAGCGCGGCGGCGGTGGATCAGGATGAGAGATCATCTCA 304
QY 61 GACACCAACGAAAGTGAATAAGTCAATGAAGAGCCATCTCTAAGAAAGAGTTCTGC 120
DB 305 GACACCAACGAAAGTGAATAAGTCAATGAAGAGCCATCTCTAAGAAAGAGTTCTGC 364
QY 121 CGGTTTGTCACTTTCCAAATCCAGTACCCTGATATTGGAAATGTATAACAGGCACAG 180
DB 365 CGGTTTGTCACTTTCCAAATCCAGTACCCTGATATTGGAAATGTATAACAGGCACAG 424
QY 181 GCTTCTCTCTGACAGCAGAGAGGTGCATCTTCAAGAGATCTCCCTCACTGGAACAAG 240
DB 425 GCTTCTCTCTGACAGCAGAGAGGTGCATCTTCAAGAGATCTCCCTCACTGGAACAAG 484
QY 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTTTAGCCTTTTTTTCAGCCAGTGAT 300
DB 485 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTTTAGCCTTTTTTTCAGCCAGTGAT 544
QY 301 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGCT 360
DB 545 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGCT 604
QY 361 CGCTGTTTCTATGCGTTTCAAAATCTCATCGAAGTGTCTCACTCAGAGATGTACAGTTG 420
DB 605 CGCTGTTTCTATGCGTTTCAAAATCTCATCGAAGTGTCTCACTCAGAGATGTACAGTTG 664
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAAATGCAATTGAA 480
DB 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAAATGCAATTGAA 724

```

```

QY 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTGGGGCTTGGCGATGATAGCAGATAGAAA 540
DB 725 ACCATGCCCTATGTTAAGAAAAAAGCAGATTGGGGCTTGGCGATGATAGCAGATAGAAA 784
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCCCTTGTCTGTAGAGAGGAGTTTCTTCTCAGGA 600
DB 785 TCTACTTTTGGGAAAGAGTGGTGGCCCTTGTCTGTAGAGAGGAGTTTCTTCTCAGGA 844
QY 601 TCTTTTGTCTATATCTTGGCTAAAGAGAGAGTCTTTATGCCAGGACTCACTTTTCTCC 660
DB 845 TCTTTTGTCTATATCTTGGCTAAAGAGAGAGTCTTTATGCCAGGACTCACTTTTCTCC 904
QY 661 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTTGTCTGCTGATGTTCCAA 720
DB 905 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTTGTCTGCTGATGTTCCAA 964
QY 721 TACTTAGTAATTAAGCCTTCAGAAAGAGGTGAGGAGATCATTTGTGATGCTGTCAAA 780
DB 965 TACTTAGTAATTAAGCCTTCAGAAAGAGGTGAGGAGATCATTTGTGATGCTGTCAAA 1024
QY 781 ATTGAGCAGGAGTTTAAACAGAGAGCCTTGCCAGTTGGCCTCATTTGGAATGAATTCGATT 840
DB 1025 ATTGAGCAGGAGTTTAAACAGAGAGCCTTGCCAGTTGGCCTCATTTGGAATGAATTCGATT 1084
QY 841 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTAATTTGTGGAACTTTGGATTCTCA 900
DB 1085 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTAATTTGTGGAACTTTGGATTCTCA 1144
QY 901 AAGGTTTTTCAGGCGAGAAATCCTTTGATTTTATGAAAAACATTTCTTTAGAGAAAA 960
DB 1145 AAGGTTTTTCAGGCGAGAAATCCTTTGATTTTATGAAAAACATTTCTTTAGAGAAAA 1204
QY 961 ACAAATTTCTTTGAGAAACGAGTTTCAGAGATATCAGCGTTTTTGAGTTATGCGAGAAACC 1020
DB 1205 ACAAATTTCTTTGAGAAACGAGTTTCAGAGATATCAGCGTTTTTGAGTTATGCGAGAAACC 1264
QY 1021 ACAGATAACGTCCTTCAACCTTGGATGCGATTTT 1053
DB 1265 ACAGATAACGTCCTTCAACCTTGGATGCGATTTT 1297

RESULT 9
AB036063
LOCUS
DEFINITION Homo sapiens p53R2 mRNA for ribonucleotide reductase, complete cds.
ACCESSION AB036063
VERSION AB036063.1 GI:7229085
KEYWORDS p53R2; ribonucleotide reductase.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (sites)
Tanaka, H., Arakawa, H., Yamaguchi, T., Shiraiishi, K., Fukuda, S.,
Matsui, K., Takei, Y. and Nakamura, Y.
A ribonucleotide reductase gene involved in a p53-dependent
cell-cycle checkpoint for DNA damage
Nature 404 (6773), 42-49 (2000)
JOURNAL MEDLINE 20179179
PUBMED 10716435
REFERENCE 2 (bases 1 to 4955)
AUTHORS Tanaka, H., Arakawa, H. and Nakamura, Y.
TITLE Direct Submission
JOURNAL Submitted (18-DEC-1999) Yusuke Nakamura, University of Tokyo,
Institute of Medical Science, Human Genome Center, Laboratory of
Molecular Medicine; 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639,
Japan (E-mail: yusuke@ims.u-tokyo.ac.jp, Tel.+81-3-5449-5372,
Fax:+81-3-5449-5433)
FEATURES
source
1..4955
/organism="Homo sapiens"
/mol_type="mRNA"

```

```

/db_xref="taxon:9606"
/chromosome="8"
/map="8q23.1"
1..4955
/gene="p53R2"
245..1300
/gene="p53R2"
/codon_start=1
/product="ribonucleotide reductase"
/protein_id="BA92434.1"
/db_xref="GI:7229086"
/translation="WGDPERPEAAGLDODERSSDTNSESFIKSNBEPILRKSRREVI
FPIQPDIMWMTKQASFTAEEDVLSKDLPHNKLKADEKFIHILAPFAASDGI
VNNELVERESQVQPEARCFYQILINVSHEMYSLIDITYIDRPKKRELFNLAIE
TMPYVKADWALRIADRSKTFGERVAFVAVEGFFSGFAAI PMLKKEFLPGLIT
FNNELISRDEGLHCPACILMFOVLNKPSEBRVREIIVDAVKIEQBFLEALPVGILG
MNCILMKOYIEBVDRLIVELGFSKVFQAEINPDMENISLEKTNFPPEKRVSEYQRF
AVMAETDINFTLDADF"
4933
polya_site
/gene="p53R2"
/note="22 a nucleotides"

ORIGIN
Query Match 100.0%; Score 1053; DB 9; Length 4955;
Best Local Similarity 100.0%; Pred. No. 1.le-255;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGCGCCGGAAGCGCGGAGCGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
DB 245 ATGGCGGACCCGGAAGCGCGGAGCGCGGCGCTGGATCAGGATGAGAGATCATCTTCA 304
QY 61 GACACCAACGAAAGTGAATAAGTCAAAATGAAGAGCCATCTCTTAAGAAAGAGTTCTCGC 120
DB 305 GACACCAACGAAAGTGAATAAGTCAAAATGAAGAGCCATCTCTTAAGAAAGAGTTCTCGC 364
QY 121 CGGTTTGTGATCTTCAATCCAGTACCTCTGATATTTGGAAATGATATAACAGGCACAG 180
DB 365 CGGTTTGTGATCTTCAATCCAGTACCTCTGATATTTGGAAATGATATAACAGGCACAG 424
QY 181 GCTTCTCTTGGACAGCAGAGAGTGCAGTATCAAGAGATCTCCTCACTGGAACAAG 240
DB 425 GCTTCTCTTGGACAGCAGAGAGTGCAGTATCAAGAGATCTCCTCACTGGAACAAG 484
QY 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCAATCTTACATCTTGGAGCGGCTTTTTCAGCGCAGTGAT 300
DB 485 CTTAAAGCAGATGAGAAGTACTTCACTCTCAATCTTACATCTTGGAGCGGCTTTTTCAGCGCAGTGAT 544
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
DB 545 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 604
QY 361 CGCTGTTTCTATGCGTTCATATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTG 420
DB 605 CGCTGTTTCTATGCGTTCATATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTG 664
QY 421 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATGAA 480
DB 665 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATGAA 724
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCGCTTGGATGGATAGCAGATAGAAAA 540
DB 725 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCGCTTGGATGGATAGCAGATAGAAAA 784
QY 541 TCTACTTTGGGAAAAGAGTGGGCTTTGCTGCTGTAGAGAGTTCCTTCTCAGGA 600
DB 785 TCTACTTTGGGAAAAGAGTGGGCTTTGCTGCTGTAGAGAGTTCCTTCTCAGGA 844
QY 601 TCTTTTGTGCTATATCTGGCTAAAGACAGAGAGTCTTATGCCAGGACTCACCTTTTTC 660
DB 845 TCTTTTGTGCTATATCTGGCTAAAGACAGAGAGTCTTATGCCAGGACTCACCTTTTTC 904
QY 661 AATGAACATCATCAGCAGATGAAGGACTTCACTGTGACATTTGCTGCGCTGATGTTCCAA 720

```

```

DB 905 AATGAACATCATCAGCAGATGAAGGACTTCACTGTGACTTTGCTGCTGATGTTCCAA 964
QY 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 780
DB 965 TACTTAGTAAATAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 1024
QY 781 ATTGAGCAGGAGTTTTTAACAGAAAGCCTTCCAGTTCGCTCATTTGGAATGAATGCAAT 840
DB 1025 ATTGAGCAGGAGTTTTTAACAGAAAGCCTTCCAGTTCGCTCATTTGGAATGAATGCAAT 1084
QY 841 TTGATGAAAACAGTACATTTAGCTTGTAGCTGACAGATTAATCTTTGGAACCTTGGATTTCTCA 900
DB 1085 TTGATGAAAACAGTACATTTAGCTTGTAGCTGACAGATTAATCTTTGGAACCTTGGATTTCTCA 1144
QY 901 AAGCTTTTTCAGCAGAAAAATCCTTTTGAATTTATGAAACAAATCTTTTGAAGAGAAAA 960
DB 1145 AAGCTTTTTCAGCAGAAAAATCCTTTTGAATTTATGAAACAAATCTTTTGAAGAGAAAA 1204
QY 961 ACAATTTTCTTTGAGAAACAGATTTTCAAGATATCAGCGTTTTCAGATTTATGCGAGAAACC 1020
DB 1205 ACAATTTTCTTTGAGAAACAGATTTTCAAGATATCAGCGTTTTCAGATTTATGCGAGAAACC 1264
QY 1021 ACAGATAACGCTTTCACCTTTGGATGCAATTTT 1053
DB 1265 ACAGATAACGCTTTCACCTTTGGATGCAATTTT 1297

RESULT 10
AR454877
LOCUS AR454877 1053 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 12 from patent US 6682917.
ACCESSION AR454877
VERSION AR454877.1 GI:42688832
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1053)
AUTHORS Nakamura,Y., Arakawa,H. and Tanaka,H.
TITLE Protein having a ribonucleotide Reductase activity and a DNA
theresof
JOURNAL Patent: US 6682917-A 12 27-JAN-2004;
FEATURES
source
1..1053
/mol_type="genomic DNA"

ORIGIN
Query Match 99.8%; Score 1051.4; DB 6; Length 1053;
Best Local Similarity 99.9%; Pred. No. 3e-255;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGCGCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
DB 1 ATGGCGGACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
QY 61 GACACCAACGAAAGTGAATAAGTCAAAATGAAGAGCCATCTCTTAAGAAAGAGTTCTCGC 120
DB 61 GACACCAACGAAAGTGAATAAGTCAAAATGAAGAGCCATCTCTTAAGAAAGAGTTCTCGC 120
QY 121 CGGTTTGTGATCTTTCATCCAGTACCTGATATTTGGAAAAATGATATAACAGGCACAG 180
DB 121 CGGTTTGTGATCTTTCATCCAGTACCTGATATTTGGAAAAATGATATAACAGGCACAG 180
QY 181 GCTTCTCTTGGACAGCAGAGAGTGCAGTATCAAGAGATCTCCCTCACTGGAACAAG 240
DB 181 GCTTCTCTTGGACAGCAGAGAGTGCAGTATCAAGAGATCTCCCTCACTGGAACAAG 240
QY 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCAGATCTTTAGCCTTTTTCAGCGCAGTGAT 300
DB 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCAGATCTTTAGCCTTTTTCAGCGCAGTGAT 300
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360

```



```
Db 301 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGAGTGACGGTTCAGAGGCT 360
Qy 361 CGCTGTTTCTATGGCTTCAAAATTCATCAGAAATGTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGGCTTCAAAATTCATCAGAAATGTTCACTCAGAGATGTACAGTTTG 420
Qy 421 CTGATAGACACTTACATCAGAGATGCCAAGAAAAGGAATTTTATTAATGCAATGAA 480
Db 421 CTGATAGACACTTACATCAGAGATGCCAAGAAAAGGAATTTTATTAATGCAATGAA 480
Qy 481 ACCATGCCCTATCTTAAGAAAAGAGAGATTTGGCCCTTCGATGGATAGCAGATAGAAA 540
Db 481 ACCATGCCCTATCTTAAGAAAAGAGAGATTTGGCCCTTCGATGGATAGCAGATAGAAA 540
Qy 541 TCTACTTTTGGGAAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGTCTTCTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGTCTTCTCTCAGGA 600
Qy 601 TCTTTTGTCTATATTTGGCTTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
Db 601 TCTTTTGTCTATATTTGGCTTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
Qy 661 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 661 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Qy 721 TACTTAGTAAATAAGCCTTCAGAAAGAGAGTCAAGGTCAGGAGATCAATGTTGATGCTGTCAA 780
Db 721 TACTTAGTAAATAAGCCTTCAGAAAGAGAGTCAAGGTCAGGAGATCAATGTTGATGCTGTCAA 780
Qy 781 ATTGAGCAGGAGTTTTTAAAGAGGCTTGGCAGTTGGCTCAATGGAATGAATTCAT 840
Db 781 ATTGAGCAGGAGTTTTTAAAGAGGCTTGGCAGTTGGCTCAATGGAATGAATTCAT 840
Qy 841 TTGATGAAAACAGTACATGAGTTGTAGCTGACAGATTAATGTTGAACTTGAATTC 900
Db 841 TTGATGAAAACAGTACATGAGTTGTAGCTGACAGATTAATGTTGAACTTGAATTC 900
Qy 901 AAGTTTTTTCAGCAGAGAAATCCTTTTGAATTTATGAAAACATTTCTTTAGAGGAAA 960
Db 901 AAGTTTTTTCAGCAGAGAAATCCTTTTGAATTTATGAAAACATTTCTTTAGAGGAAA 960
Qy 961 ACAAATTTCTTTCAGAAAACAGTTTCAGAGTATCAGCGTTTTCAGATTATGGCAGAAACC 1020
Db 961 ACAAATTTCTTTCAGAAAACAGTTTCAGAGTATCAGCGTTTTCAGATTATGGCAGAAACC 1020
Qy 1021 ACAGATAACGCTTCTACCTTTGGATGAGATTTT 1053
Db 1021 ACAGATAACGCTTCTACCTTTGGATGAGATTTT 1053

RESULT 11
BD064773
LOCUS Novel protein and DNA thereof.
DEFINITION Novel protein and DNA thereof.
ACCESSION BD064773
VERSION BD064773.1 GI:22610376
KEYWORDS JP 2001269184-A/10.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1053)
Nakamura, Y., Arakawa, H. and Tanaka, H.
Novel protein and DNA thereof
Patent: JP 2001269184-A 10 02-OCT-2001.
YUSUKE NAKAMURA, TAKEDA CHEMICAL INDUSTRIES LTD
OS Homo sapiens (human)
PN JP 2001269184-A/10
PD 02-OCT-2001
PF 27-JUN-2000 JP 2000192401
PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
```

```
FEATURES
source
1..1053
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 99.8%; Score 1051.4; DB 6; Length 1053;
Best Local Similarity 99.9%; Pred. No. 3e-255;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGGCGCAGCCCGGAAAGCGCGGAGCGGCTGGATCAGAGATGAGAGATCATCTTCA 60
Db 1 ATGGCGCAGCCCGGAAAGCGCGGAGCGGCTGGATCAGAGATGAGAGATCATCTTCA 60
Qy 61 GACACCAACCAAGTGAATTAAGTCAANTGAAGGCCACTCTTAAGAAAGAGTCTCGC 120
Db 61 GACACCAACCAAGTGAATTAAGTCAANTGAAGGCCACTCTTAAGAAAGAGTCTCGC 120
Qy 121 CGGTTTCTCATCTTCCAAATCCAGTACCTCATATTTGGAAAATGTATAAACAGGCACAG 180
Db 121 CGGTTTCTCATCTTCCAAATCCAGTACCTCATATTTGGAAAATGTATAAACAGGCACAG 180
Qy 181 GCTTCTCTTGGACAGAGAGAGTCCGACTTATCAAGAGATCTCCCTCACTGGAAACAAG 240
Db 181 GCTTCTCTTGGACAGAGAGAGTCCGACTTATCAAGAGATCTCCCTCACTGGAAACAAG 240
Qy 241 CTTAAAGCAGATGAGAGTACTTCTCATCTCAGATCTTAGCGCTTTTTCGACGCGAGTAT 300
Db 241 CTTAAAGCAGATGAGAGTACTTCTCATCTCAGATCTTAGCGCTTTTTCGACGCGAGTAT 300
Qy 301 GGAATTCGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGAGTGCAGGTTCCAGAGCT 360
Db 301 GGAATTCGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGAGTGCAGGTTCCAGAGCT 360
Qy 361 CGCTGTTTCTATGGCTTTCAAAATTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGGCTTTCAAAATTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 420
Qy 421 CTGATAGACACTTACATCAGAGATGCCAAGAAAAGGAATTTTATTAATGCAATGAA 480
Db 421 CTGATAGACACTTACATCAGAGATGCCAAGAAAAGGAATTTTATTAATGCAATGAA 480
Qy 481 ACCATGCCCTATCTTAAGAAAAGAGAGTGGCCCTTCGATGGATAGCAGATAGAAA 540
Db 481 ACCATGCCCTATCTTAAGAAAAGAGAGTGGCCCTTCGATGGATAGCAGATAGAAA 540
Qy 541 TCTACTTTTGGGAAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGTCTTCTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGTCTTCTCTCAGGA 600
Qy 601 TCTTTTGTCTATATTTGGCTTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
Db 601 TCTTTTGTCTATATTTGGCTTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
Qy 661 AATGAACCTCATCAGCAGATGAAGAGTTCACCTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 661 AATGAACCTCATCAGCAGATGAAGAGTTCACCTGTGACTTTGCTTGCCTGATGTTCCAA 720
Qy 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGTCAGGAGATCAATGTTGATGCTGTCAA 780
```



```
|||||TACTAGTAATAAGCCTTCAGAAAGAGGCTCAGGAGATCAATGTTGATGCTGTCAA 780
Db
781 ATTAGCAGAGAGTTTAAACAGAGAGCTTCCAGTTGGCCTCAATGGAATGAATTCAT 840
Qy
781 ATTAGCAGAGAGTTTAAACAGAGAGCTTCCAGTTGGCCTCAATGGAATGAATTCAT 840
Db
841 TTGATGAACACGATACATGAGTTGTTAGCTGACAGATTAATCTGTGAACTTGGATTCTCA 900
Qy
841 TTGATGAACACGATACATGAGTTGTTAGCTGACAGATTAATCTGTGAACTTGGATTCTCA 900
Db
901 AAGGTTTTTCNGCAGCAAAATCCTTTTGATTTTATGGAACCAATTTCTTTAGAGGAAA 960
Qy
901 AAGGTTTTTCNGCAGCAAAATCCTTTTGATTTTATGGAACCAATTTCTTTAGAGGAAA 960
Db
961 ACAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGAGTTATGGCAGAAACC 1020
Qy
961 ACAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGAGTTATGGCAGAAACC 1020
Db
1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053
Qy
1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053
Db

RESULT 12
BD093086
LOCUS
DEFINITION Novel protein and its DNA.
ACCESSION BD093086
VERSION BD093086.1 GI:22638674
KEYWORDS WO 0100799-A/10.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1053)
Nakamura,Y., Arakawa,H. and Tanaka.H.
Novel protein and its DNA
Patent: WO 0100799-A 10 04-JAN-2001;
TAKEDA CHEMICAL INDUSTRIES LTD,YUSUKE NAKAMURA,HIROFUMI ARAKAWA,
HIROSHI TANAKA
OS Homo sapiens (human)
PN WO 0100799-A/10
PD 04-JAN-2001
PF 27-JUN-2000 WO 2000JP004189
PR 28-JUN-1999 JP 99P 181131,06-JUL-1999 JP 99P 192391 PR
21-JAN-2000 JP 00P 017770
PI YUSUKE NAKAMURA,HIROFUMI ARAKAWA,HIROSHI TANAKA PC
C12N9/04,C12N15/53,C12N1/15,C12N1/19,C12N1/21,C12N5/10 PC
C12P21/02,A61K38/44,
PC A61K48/00,C07K16/40,G01N33/53,C12Q1/26
CC
FH Key Location/Qualifiers.
FEATURES
source
1..1053
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 99.8%; Score 1051.4; DB 6; Length 1053;
Best Local Similarity 99.9%; Pred. No. 3e-255;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGGGGACCCGAAAGCGCGGAGCGCGGCTGGATCAGATGAGATCAATCTTCA 60
Db 1 ATGGGGACCCGAAAGCGCGGAGCGCGGCTGGATCAGATGAGATCAATCTTCA 60
Qy 61 GACACCAACCAAGTGAATAAGTCAATGAGAGCCACTCTTGAAGAGTTCTCGC 120
Db 61 GACACCAACCAAGTGAATAAGTCAATGAGAGCCACTCTTGAAGAGTTCTCGC 120
Qy 121 CGGTTTGTCTATCTTCCAAATCCAGTACCCTGATATTTGGAAATGTATAAACAGGCACAG 180
```

```
|||||CGGTTTGTCTATCTTCCATCCAGTACCCTGATATTTGGAAATGTATAAACAGGCACAG 180
Db
181 GCTTCTCTTCTGGACAGCAGAGAGGTCGACTTATCAAAGGATCTCCCTCACTGGAACAAG 240
Qy
181 GCTTCTCTTCTGGACAGCAGAGAGGTCGACTTATCAAAGGATCTCCCTCACTGGAACAAG 240
Db
241 CTTAAAGCAGATGAGAAAGTACTTCTCATCTCTCAATCTTAGCCCTTTTTCGAGCCAGTGAT 300
Qy
241 CTTAAAGCAGATGAGAAAGTACTTCTCATCTCTCAATCTTAGCCCTTTTTCGAGCCAGTGAT 300
Db
301 GGAATTCGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGGCT 360
Qy
301 GGAATTCGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGGCT 360
Db
361 CGCTGTTTCTATGCTTTCAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 420
Qy
361 CGCTGTTTCTATGCTTTCAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 420
Db
421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTGAA 480
Qy
421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTGAA 480
Db
481 ACCATGCCCTATGTTAAGAAAGAGATTTGGGCCCTTGGATGGATAGCAGATAGAAA 540
Qy
481 ACCATGCCCTATGTTAAGAAAGAGATTTGGGCCCTTGGATGGATAGCAGATAGAAA 540
Db
541 TCTACTTTTGGGAAAGAGTGGTGGCCCTTCTCTCTAGAGAGAGTTTCTTCTCAGGA 600
Qy
541 TCTACTTTTGGGAAAGAGTGGTGGCCCTTCTCTCTAGAGAGAGTTTCTTCTCAGGA 600
Db
601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGTCTTATGCCAGGACTCACATTTTTC 660
Qy
601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGTCTTATGCCAGGACTCACATTTTTC 660
Db
661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Qy
661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db
721 TACTTAGTAATAAGCCTTCAGAAAGAGGCTCAGGAGATCAATGTTGATGCTGTCAAA 780
Qy
721 TACTTAGTAATAAGCCTTCAGAAAGAGGCTCAGGAGATCAATGTTGATGCTGTCAAA 780
Db
781 ATTGAGCAGGAGTTTAAACAGAGAGCTTGCAGTTGGCCTCATTTGGAATGAATTCAT 840
Qy
781 ATTGAGCAGGAGTTTAAACAGAGAGCTTGCAGTTGGCCTCATTTGGAATGAATTCAT 840
Db
841 TTGATGAACACGATACATGAGTTTGTAGCTGACAGATTAATCTTGGAACTTGGATTCTCA 900
Qy
841 TTGATGAACACGATACATGAGTTTGTAGCTGACAGATTAATCTTGGAACTTGGATTCTCA 900
Db
901 AAGGTTTTTCAGGAGCAAAATCCTTTTGATTTTATGGAACCAATTTCTTTAGAGGAAA 960
Qy
901 AAGGTTTTTCAGGAGCAAAATCCTTTTGATTTTATGGAACCAATTTCTTTAGAGGAAA 960
Db
961 ACAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGAGTTATGGCAGAAACC 1020
Qy
961 ACAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGAGTTATGGCAGAAACC 1020
Db
1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053
Qy
1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053
Db

RESULT 13
AR454870
LOCUS
DEFINITION Sequence 4 from patent US 6682917.
ACCESSION AR454870
VERSION AR454870.1 GI:42688825
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
```

Unclassified.
1 (bases 1 to 1081)
Nakamura, Y., Arakawa, H. and Tanaka, H.
Protein having a ribonucleotide Reductase activity and a DNA
thereof
Patent: US 682917-A 4 27-JAN-2004;
JOURNAL
FEATURES
source Location/Qualifiers
1. .1081
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 99.8%; Score 1051.4; DB 6; Length 1081;
Best Local Similarity 99.9%; Pred. No. 3e-255;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCGGAAAGCGCGGAGCGGCTGGATCAGATGAGAGATCATCTTCA 60
DB 20 ATGGGCGACCGGAAAGCGGAGCGGCTGGATCAGATGAGAGATCATCTTCA 79
QY 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
DB 80 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 139
QY 121 CGGTTGTCTATCTTCCAAATCCAGTACCTGTATTTGGAAATGTATAAACAGGCACAG 180
DB 140 CGGTTGTCTATCTTCCAAATCCAGTACCTGTATTTGGAAATGTATAAACAGGCACAG 199
QY 181 GCTTCCTCTGGACAGCAGAGAGTGCATTTATCAAGAGATCTCCCTCACTGGAACAAG 240
DB 200 GCTTCCTCTGGACAGCAGAGAGTGCATTTATCAAGAGATCTCCCTCACTGGAACAAG 259
QY 241 CTTTAAAGCAGATGAGAGTACTTCTCATCTCTACATCTTAGCTTTTGGCAGCAGTGAT 300
DB 260 CTTTAAAGCAGATGAGAGTACTTCTCATCTCTACATCTTAGCTTTTGGCAGCAGTGAT 319
QY 301 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCGAGTTCCAGAGGCT 360
DB 320 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCGAGTTCCAGAGGCT 379
QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAAGATGTTCACTCAGAGATGTACAGTTG 420
DB 380 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAAGATGTTCACTCAGAGATGTACAGTTG 439
QY 421 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATGAA 480
DB 440 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATGAA 499
QY 481 ACCATGCCCTATGTTAAGAAAGAGAGATTTGGGCGCTTGGATGAGATAGAGAA 540
DB 500 ACCATGCCCTATGTTAAGAAAGAGAGATTTGGGCGCTTGGATGAGATAGAGAA 559
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAAGAGTTTCTTCTCAGGA 600
DB 560 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAAGAGTTTCTTCTCAGGA 619
QY 601 TCTTTTGTGCTATATCTGCTTAAGAAAGAGAGTCTTATGCCAGACTCATTCTTC 660
DB 620 TCTTTTGTGCTATATCTGCTTAAGAAAGAGAGTCTTATGCCAGACTCATTCTTC 679
QY 661 AATGAATCATCAGCAGATGAAGACTTCTCATGTGACTTTTGTCTGCTGATGTTCCAA 720
DB 680 AATGAATCATCAGCAGATGAAGACTTCTCATGTGACTTTTGTCTGCTGATGTTCCAA 739
QY 721 TACTTAGTAAATAGCCCTTCAGAAAGAGGCTCAGGAGATCATTTGTGTGCTGTCAAA 780
DB 740 TACTTAGTAAATAGCCCTTCAGAAAGAGGCTCAGGAGATCATTTGTGTGCTGTCAAA 799
QY 781 ATTGAGCAGAGTTTTTAAACAGAGCCTTGGCAGTTGGCCCTCATTTGAATGAAATTCATT 840
DB 800 ATTGAGCAGAGTTTTTAAACAGAGCCTTGGCAGTTGGCCCTCATTTGAATGAAATTCATT 859
QY 841 TTGATGAAACAGTACATTTGAGTTTGTAGCTGACAGATTACTTGTGGAACTTGGATTCTCA 900

Db 860 TTGATGAAACAGTACATTTGAGTTTGTAGCTGACAGATTACTTGTGGAACTTGGATTCTCA 919
QY 901 AAGGTTTTTTCAGGCGAGAAATCCTTTTTCATTTTATGGAACAACTTTCTTTAGAGGAAA 960
DB 920 AAGGTTTTTTCAGGCGAGAAATCCTTTTTCATTTTATGGAACAACTTTCTTTAGAGGAAA 979
QY 961 ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTTTCAGTTATGCGCAGAAACC 1020
DB 980 ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTTTCAGTTATGCGCAGAAACC 1039
QY 1021 ACAGATAACGCTTTCACCTTGGATGCGAGATTTT 1053
DB 1040 ACAGATAACGCTTTCACCTTGGATGCGAGATTTT 1072

RESULT 14
BD064766 1081 bp DNA linear PAT 27-AUG-2002
LOCUS Novel protein and DNA thereof.
DEFINITION BD064766
ACCESSION BD064766
VERSION BD064766.1 GI:22610369
KEYWORDS JP 2001269184-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1081)
AUTHORS Nakamura, Y., Arakawa, H. and Tanaka, H.
TITLE Novel protein and DNA thereof
JOURNAL YUSUKE NAKAMURA, TAKEDA CHEMICAL INDUSTRIES LTD
COMMENT OS Homo sapiens (human)
PN JP 2001269184-A/3
PD 02-OCT-2001
PF 27-JUN-2000 JP 2000192401
PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
A61P35/00, A61K31/711, A61K38/00, A61K38/43, A61K45/00, A61K48/00, PC
A61P43/00, C07K16/32, C07K16/40, C12N1/15, C12N1/19, C12N1/21 PC
C12N5/10, C12N9/02,
PC C12P21/02, C12P21/08, G01N33/15, G01N33/50, G01N33/53// (C12N15/09,
PC C12R1:91),
PC (C12N15/09, C12R1:19), (C12N9/02, C12R1:91), (C12P21/02, C12R1:91),
PC (C12P21/08, C12R1:91), C12N15/00, A61K37/02, A61K37/48, C12N5/00,
PC (C12N15/00, C12R1:91), (C12N15/00, C12R1:19)
CC Novel protein and DNA thereof
FH Key Location/Qualifiers
FT source 1. .1081
FT Location/Qualifiers
1. .1081
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES
source
1. .1081
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 99.8%; Score 1051.4; DB 6; Length 1081;
Best Local Similarity 99.9%; Pred. No. 3e-255;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCGGAAAGCGCGGAGCGGCTGGATCAGATGAGAGATCATCTTCA 60
DB 20 ATGGGCGACCGGAAAGCGGAGCGGCTGGATCAGATGAGAGATCATCTTCA 79
QY 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
DB 80 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 139
QY 121 CGGTTGTCTATCTTCCAAATCCAGTACCTGTATTTGGAAATGTATAAACAGGCACAG 180
DB 140 CGGTTGTCTATCTTCCAAATCCAGTACCTGTATTTGGAAATGTATAAACAGGCACAG 199
QY 181 GCTTCCTCTGGACAGCAGAGAGTGCATTTATCAAGAGATCTCCCTCACTGGAACAAG 240
DB 200 GCTTCCTCTGGACAGCAGAGAGTGCATTTATCAAGAGATCTCCCTCACTGGAACAAG 259
QY 241 CTTTAAAGCAGATGAGAGTACTTCTCATCTCTACATCTTAGCTTTTGGCAGCAGTGAT 300
DB 260 CTTTAAAGCAGATGAGAGTACTTCTCATCTCTACATCTTAGCTTTTGGCAGCAGTGAT 319
QY 301 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCGAGTTCCAGAGGCT 360
DB 320 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCGAGTTCCAGAGGCT 379
QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAAGATGTTCACTCAGAGATGTACAGTTG 420
DB 380 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAAGATGTTCACTCAGAGATGTACAGTTG 439
QY 421 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATGAA 480
DB 440 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATGAA 499
QY 481 ACCATGCCCTATGTTAAGAAAGAGAGATTTGGGCGCTTGGATGAGATAGAGAA 540
DB 500 ACCATGCCCTATGTTAAGAAAGAGAGATTTGGGCGCTTGGATGAGATAGAGAA 559
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAAGAGTTTCTTCTCAGGA 600
DB 560 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAAGAGTTTCTTCTCAGGA 619
QY 601 TCTTTTGTGCTATATCTGCTTAAGAAAGAGAGTCTTATGCCAGACTCATTCTTC 660
DB 620 TCTTTTGTGCTATATCTGCTTAAGAAAGAGAGTCTTATGCCAGACTCATTCTTC 679
QY 661 AATGAATCATCAGCAGATGAAGACTTCTCATGTGACTTTTGTCTGCTGATGTTCCAA 720
DB 680 AATGAATCATCAGCAGATGAAGACTTCTCATGTGACTTTTGTCTGCTGATGTTCCAA 739
QY 721 TACTTAGTAAATAGCCCTTCAGAAAGAGGCTCAGGAGATCATTTGTGTGCTGTCAAA 780
DB 740 TACTTAGTAAATAGCCCTTCAGAAAGAGGCTCAGGAGATCATTTGTGTGCTGTCAAA 799
QY 781 ATTGAGCAGAGTTTTTAAACAGAGCCTTGGCAGTTGGCCCTCATTTGAATGAAATTCATT 840
DB 800 ATTGAGCAGAGTTTTTAAACAGAGCCTTGGCAGTTGGCCCTCATTTGAATGAAATTCATT 859
QY 841 TTGATGAAACAGTACATTTGAGTTTGTAGCTGACAGATTACTTGTGGAACTTGGATTCTCA 900

```
QY 181 GCTTCCTTCTGACAGCAGAGGTCGACTTATCAAGGATCTCCCTCACTGGAACAAG 240
Db 200 GCTTCCTTCTGACAGCAGAGGTTGACTTATCAAGGATCTCCCTCACTGGAACAAG 259
QY 241 CTTAAAGCAGATGAGAAGTACTCTCTCATCTCTAGCCCTTTTTCAGCCAGTGAT 300
Db 260 CTTAAAGCAGATGAGAAGTACTCTCTCATCTCTAGCCCTTTTTCAGCCAGTGAT 319
QY 301 GGAATTGTAATGAAATTTGGTGGAGCGCTTTTAGTCAGGAGGTCCAGGTTCCAGAGGCT 360
Db 320 GGAATTGTAATGAAATTTGGTGGAGCGCTTTTAGTCAGGAGGTCCAGGTTCCAGAGGCT 379
QY 361 CGCTGTTCTATGCTTCCAAATTTCTATCGAGAAATGTCATCAGAGATGACAGTTTG 420
Db 380 CGCTGTTCTATGCTTCCAAATTTCTATCGAGAAATGTCATCAGAGATGACAGTTTG 439
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 480
Db 440 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 499
QY 481 ACCATGCCCTATGTTAAGAAAGAGAGATTTGGCCCTTCGATGGATAGCAGATAGAAA 540
Db 500 ACCATGCCCTATGTTAAGAAAGAGAGATTTGGCCCTTCGATGGATAGCAGATAGAAA 559
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAAGAGAGTTCCTTCTCAGGA 600
Db 560 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAAGAGAGTTCCTTCTCAGGA 619
QY 601 TCTTTTGTCTATATTTCTGGCTTAAGAGAGAGTTCCTTCCAGGACTCACTTTTCC 660
Db 620 TCTTTTGTCTATATTTCTGGCTTAAGAGAGAGTTCCTTCCAGGACTCACTTTTCC 679
QY 661 AATGAACCTCATCAGCAGAGATCAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 680 AATGAACCTCATCAGCAGAGATCAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 739
QY 721 TACTTAGTAAATAGCTTTCAAGAAAGGGTCAGGAGATCAATGTTGATGCTGTCAAA 780
Db 740 TACTTAGTAAATAGCTTTCAAGAAAGGGTCAGGAGATCAATGTTGATGCTGTCAAA 799
QY 781 ATTGAGCAGAGTTTTAAACAGAGCCTTGCAGTTGGCTCTATTGGAATGATTCATT 840
Db 800 ATTGAGCAGAGTTTTAAACAGAGCCTTGCAGTTGGCTCTATTGGAATGATTCATT 859
QY 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTTGGAATCTCA 900
Db 860 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTTGGAATCTCA 919
QY 901 AAGGTTTTTTCAGGAGAAATCCTTTGATTTTATGGAAGAAATTTCTTTAGAGGAAAA 960
Db 920 AAGGTTTTTTCAGGAGAAATCCTTTGATTTTATGGAAGAAATTTCTTTAGAGGAAAA 979
QY 961 ACAATTTCTTTCAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC 1020
Db 980 ACAATTTCTTTCAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC 1039
QY 1021 ACAGATAAGCTCTTCACTTGGATGAGATTTT 1053
Db 1040 ACAGATAAGCTCTTCACTTGGATGAGATTTT 1072

RESULT 15
BD093079
LOCUS BD093079
DEFINITION Novel protein and its DNA.
ACCESSION BD093079
VERSION BD093079.1 GI:22638667
KEYWORDS WO 0100799-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1081)
```

```
AUTHORS Nakamura, Y., Arakawa, H. and Tanaka, H.
TITLE Novel protein and its DNA
JOURNAL Patent: WO 0100799-A 3 04-JAN-2001;
TAKEDA CHEMICAL INDUSTRIES LTD, YUSUKE NAKAMURA, HIROFUMI ARAKAWA,
HIROSHI TANAKA
COMMENT OS Homo sapiens (human)
PN WO 0100799-A/3
ED 04-JAN-2001
PF 27-JUN-2000 WO 2000JP004189
PR 28-JUN-1999 JP 99P 181131, 06-JUL-1999 JP 99P 192391 PR
21-JAN-2000 JP 00P 017770
PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
C12N9/04, C12N15/53, C12N1/15, C12N1/19, C12N1/21, C12N5/10 PC
, C12P21/02, A61K38/44,
PC A61K48/00, C07K16/40, G01N33/53, C12Q1/26
CC
FH Key Location/Qualifiers
1..1081
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 99.8%; Score 1051.4; DB 6; Length 1081;
Best Local Similarity 99.9%; Pred. No. 3e-255;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGCGGACCCGGAAGGCGGAGCGGCTGATCAGGATGAGAGATCATCTTCA 60
Db 20 ATGGCGGACCCGGAAGGCGGAGCGGCTGATCAGGATGAGAGATCATCTTCA 79
QY 61 GACACCAACGAAAGTCAAAATGAAAGTCAAAATGAAAGGACCTCTTAAGAAAGAGTTCTCGC 120
Db 80 GACACCAACGAAAGTCAAAATGAAAGTCAAAATGAAAGGACCTCTTAAGAAAGAGTTCTCGC 139
QY 121 CGTTTGTCTATCTTCCAAATCCAGTACCTGATATTTGGAAATGTAATAACAGCACAG 180
Db 140 CGTTTGTCTATCTTCCAAATCCAGTACCTGATATTTGGAAATGTAATAACAGCACAG 199
QY 181 GCTTCCTTCTGGACAGCAGAGAGGTGCGATTCAAGGATCTCCCTCACTGGAACAAG 240
Db 200 GCTTCCTTCTGGACAGCAGAGAGGTGCGATTCAAGGATCTCCCTCACTGGAACAAG 259
QY 241 CTTAAAGCAGATGAGAAGTACTTCTCTCATCTTCTAGCCCTTTTTCAGCCAGTGAT 300
Db 260 CTTAAAGCAGATGAGAAGTACTTCTCTCATCTCTAGCCCTTTTTCAGCCAGTGAT 319
QY 301 GGAATTGTAATGAAATTTGGTGGAGCGCTTTTAGTCAGGAGGTCCAGGTTCCAGAGGCT 360
Db 320 GGAATTGTAATGAAATTTGGTGGAGCGCTTTTAGTCAGGAGGTCCAGGTTCCAGAGGCT 379
QY 361 CGCTGTTCTATGCTTCCAAATTTCTATCGAGAAATGTCATCAGAGATGACAGTTTG 420
Db 380 CGCTGTTCTATGCTTCCAAATTTCTATCGAGAAATGTCATCAGAGATGACAGTTTG 439
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 480
Db 440 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 499
QY 481 ACCATGCCCTATGTTAAGAAAGAGAGATTTGGCCCTTCGATGGATAGCAGATAGAAA 540
Db 500 ACCATGCCCTATGTTAAGAAAGAGAGATTTGGCCCTTCGATGGATAGCAGATAGAAA 559
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAAGAGAGTTCCTTCTCAGGA 600
Db 560 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAAGAGAGTTCCTTCTCAGGA 619
QY 601 TCTTTTGTCTATATTTCTGGCTTAAGAGAGAGTTCCTTCCAGGACTCACTTTTCC 660
Db 620 TCTTTTGTCTATATTTCTGGCTTAAGAGAGAGTTCCTTCCAGGACTCACTTTTCC 679
QY 661 AATGAACCTCATCAGCAGAGATCAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
```

Db		680	AATGAACCTCATCAGCAGAGATGAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA	739
Qy		721	TACTTAGTAAATAAGCCTTCAGAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAAA	780
Db		740	TACTTAGTAAATAAGCCTTCAGAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAAA	799
Qy		781	ATTGACGAGAGTTTAAACAGAGCCTTCCAGTTGGCCTCATTTGGAATGAATTGCATT	840
Db		800	ATTGACGAGAGTTTAAACAGAGCCTTCCAGTTGGCCTCATTTGGAATGAATTGCATT	859
Qy		841	TTGATGAAACAGTACATTTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA	900
Db		860	TTGATGAAACAGTACATTTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA	919
Qy		901	AAGGTTTTTTCAGGCAGAAATCCTTTTGATTTTATGGAACACATTTCTTTAGAGGAAAA	960
Db		920	AAGGTTTTTTCAGGCAGAAATCCTTTTGATTTTATGGAACACATTTCTTTAGAGGAAAA	979
Qy		961	ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGGTTTATGGCAGAAAC	1020
Db		980	ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGGTTTATGGCAGAAAC	1039
Qy		1021	ACAGATAACGTTCTTACCTTGGATGAGATTTT	1053
Db		1040	ACAGATAACGTTCTTACCTTGGATGAGATTTT	1072

Search completed: June 16, 2005, 08:31:35
 Job time : 4893 secs

Result No.	Query		DB	ID	Description
	Score	Match Length			
1	1053	100.0	1053	4	US-10-019-733-2
2	1053	100.0	4955	4	US-10-019-733-3
3	1051.4	99.8	1053	4	US-10-019-733-12
4	1051.4	99.8	1081	4	US-10-019-733-4
5	579	55.0	2500	4	US-09-962-665-9
6	579	55.0	2500	4	US-09-963-333-9
7	579	55.0	2500	4	US-09-962-677-9
8	578.4	54.9	2479	4	US-09-949-016-2025
9	578.4	54.9	2500	4	US-09-023-655-1370
10	578.4	54.9	2500	4	US-09-949-016-145
11	441	41.9	481	3	US-08-905-223-125
12	429.8	40.8	14176	1	US-08-307-499-1
13	429.8	40.8	14176	1	US-08-307-499-14
14	429.8	40.8	14176	3	US-09-299-268-1
15	429.8	40.8	14176	3	US-09-299-268-14
16	383.2	36.4	1236	4	US-09-248-796A-3914
17	382.8	36.4	825	1	US-08-307-499-40
18	382.8	36.4	825	3	US-09-299-268-40
19	337.8	32.1	1112	1	US-08-136-743B-1
20	318.2	30.2	419	3	US-08-905-223-129
21	316	30.0	1083	4	US-09-248-796A-3915
22	260.8	24.8	656	3	US-09-328-111-791
23	195.2	18.5	449	4	US-09-370-767-14670
24	169.2	16.1	366	1	US-08-307-499-13
25	169.2	16.1	366	3	US-09-299-268-13
26	141.4	13.4	373	4	US-09-401-064-337
27	139.8	13.3	373	4	US-09-401-064-339
					Sequence 2, Appli
					Sequence 3, Appli
					Sequence 12, Appli
					Sequence 4, Appli
					Sequence 9, Appli
					Sequence 9, Appli
					Sequence 9, Appli
					Sequence 2025, Appli
					Sequence 1370, Appli
					Sequence 145, Appli
					Sequence 125, Appli
					Sequence 1, Appli
					Sequence 14, Appli
					Sequence 1, Appli
					Sequence 3914, Appli
					Sequence 40, Appli
					Sequence 1, Appli
					Sequence 129, Appli
					Sequence 3915, Appli
					Sequence 791, Appli
					Sequence 14670, A
					Sequence 13, Appli
					Sequence 13, Appli
					Sequence 337, Appli
					Sequence 319, Appli

Db 301 GGAATTTGTAATCAAAATTTGTTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
Qy 361 CCCTGTTTCTATGGCTTTCAAAATTTCTATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db 361 CCCTGTTTCTATGGCTTTCAAAATTTCTATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATGAA 480
Qy 481 ACCATGCCCTATGTTAAGAAAAGAGAGATTTGGCCCTTGGATGGATAGAGATAGAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAAGAGAGATTTGGCCCTTGGATGGATAGAGATAGAAA 540
Qy 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAAGAGAGTCTTCTCTCAGGA 600
Db 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAAGAGAGTCTTCTCTCAGGA 600
Qy 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGTCTTATGCGAGACTCACTTTTTC 660
Db 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGTCTTATGCGAGACTCACTTTTTC 660
Qy 661 AATGAACCTCATCAGCAGATGAAGACTTCTACTGTGACTTTGCTGCTGATGTTCCAA 720
Db 661 AATGAACCTCATCAGCAGATGAAGACTTCTACTGTGACTTTGCTGCTGATGTTCCAA 720
Qy 721 TACTTAGTAATAAGCTTTCAGAGAAAAGGCTCAGGAGAGATCACTTGTGATGCTGCAAA 780
Db 721 TACTTAGTAATAAGCTTTCAGAGAAAAGGCTCAGGAGAGATCACTTGTGATGCTGCAAA 780
Qy 781 ATTGAGCAGGAGTTTTTAAACAGAGCCTTTGCCAGTTGGCTCAATGGAATGAATTCATT 840
Db 781 ATTGAGCAGGAGTTTTTAAACAGAGCCTTTGCCAGTTGGCTCAATGGAATGAATTCATT 840
Qy 841 TTGATGAACACGTATTCAGTTTGTAGCTGACAGATTAATCTGTGGAACTTGGATTCCTCA 900
Db 841 TTGATGAACACGTATTCAGTTTGTAGCTGACAGATTAATCTGTGGAACTTGGATTCCTCA 900
Qy 901 AAGGTTTTTCCAGCAGAAAATCCTTTTGAATTTTATGGAAGAACTTTCTTTAGAGGAAA 960
Db 901 AAGGTTTTTCCAGCAGAAAATCCTTTTGAATTTTATGGAAGAACTTTCTTTAGAGGAAA 960
Qy 961 ACAAAATTTCTTTGAGAAAACGATTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC 1020
Db 961 ACAAAATTTCTTTGAGAAAACGATTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC 1020
Qy 1021 ACAGATAACGTCTTCACTTGGATGAGATTTT 1053
Db 1021 ACAGATAACGTCTTCACTTGGATGAGATTTT 1053

RESULT 2

US-10-019-733-3
; Sequence 3, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619W00P
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 3
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens

; FEATURE:
; OTHER INFORMATION:
US-10-019-733-3

Query Match 100.0%; Score 1053; DB 4; Length 4955;
Best Local Similarity 100.0%; Pred. No. 2.1e-310;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGGCGACCCGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 245 ATGGGCGACCCGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
Qy 61 GACACCAACCAAGTGAATAAAGTCAANTGAAGAGCCACTCCTTAAGAAAGATTTCTGC 120
Db 305 GACACCAACCAAGTGAATAAAGTCAANTGAAGAGCCACTCCTTAAGAAAGATTTCTGC 364
Qy 121 CGGTTTGTCTATCTTTCCAAATCCAGTACCTGATATTTGGAAAAATGTATAAAGAGCAG 180
Db 365 CGGTTTGTCTATCTTTCCAAATCCAGTACCTGATATTTGGAAAAATGTATAAAGAGCAG 424
Qy 181 GCTTCTCTTCTGGACAGCAGAGAGTGCAGTTCATCAAGAGATCTCCCTCACTGGAAACAAG 240
Db 425 GCTTCTCTTCTGGACAGCAGAGAGTGCAGTTCATCAAGAGATCTCCCTCACTGGAAACAAG 484
Qy 241 CTTAAGCAGATGAGAAAGTACTTCTCATCTCTCAGATCTTAGCCTTTTTCAGGCCAGTGAT 300
Db 485 CTTAAGCAGATGAGAAAGTACTTCTCATCTCTCAGATCTTAGCCTTTTTCAGGCCAGTGAT 544
Qy 301 GGAATTTGTAATAAGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGAGGTTCCAGAGGCT 360
Db 545 GGAATTTGTAATAAGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGAGGTTCCAGAGGCT 604
Qy 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAGAATGTTCTCACTCAGAGATGTACAGTTTG 420
Db 605 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAGAATGTTCTCACTCAGAGATGTACAGTTTG 664
Qy 421 CTGATAGACACTTATACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATGAA 480
Db 665 CTGATAGACACTTATACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATGAA 724
Qy 481 ACCATGCCCTATGTTAAGAAAAGAGAGATTTGGCCCTTGGATGGATAGAGATAGAAA 540
Db 725 ACCATGCCCTATGTTAAGAAAAGAGAGATTTGGCCCTTGGATGGATAGAGATAGAAA 784
Qy 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAAAGAGTCTTCTCTCAGGA 600
Db 785 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAAAGAGTCTTCTCTCAGGA 844
Qy 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGTTTATGCGAGACTCACTTTTTC 660
Db 845 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGTTTATGCGAGACTCACTTTTTC 904
Qy 661 AATGAACCTCATCAGCAGATGAGAGTTCACCTGATGCTTTGCTGCTGATGTTCCAA 720
Db 905 AATGAACCTCATCAGCAGATGAGAGTTCACCTGATGCTTTGCTGCTGATGTTCCAA 964
Qy 721 TACTTAGTAATAAGCTTTCAGAGAAAAGGCTCAGGAGAGATCACTTGTGATGCTGCTCAAA 780
Db 965 TACTTAGTAATAAGCTTTCAGAGAAAAGGCTCAGGAGAGATCACTTGTGATGCTGCTCAAA 1024
Qy 781 ATTGAGCAGGAGTTTTTAAACAGAGCCTTGCAGTTGCGCTCATTTGGAAATGAATTCATT 840
Db 1025 ATTGAGCAGGAGTTTTTAAACAGAGCCTTGCAGTTGCGCTCATTTGGAAATGAATTCATT 1084
Qy 841 TTGATGAACACGTATTCAGTTTGTAGCTGACAGATTAATCTGTGGAACCTTGGATTCCTCA 900
Db 1085 TTGATGAACACGTATTCAGTTTGTAGCTGACAGATTAATCTGTGGAACCTTGGATTCCTCA 1144
Qy 901 AAGGTTTTTCCAGCAGAAAATCCTTTTGAATTTTATGGAAGAACTTTCTTTAGAGGAAA 960
Db 1145 AAGGTTTTTCCAGCAGAAAATCCTTTTGAATTTTATGGAAGAACTTTCTTTAGAGGAAA 1204
Qy 961 ACAAAATTTCTTTGAGAAAACGAGTTTTCAGAGTATCAGCGCTTTTTCAGAGTTATGCGAGAAACC 1020

Db 1205 ACAAATTTCTTTCAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGAGTTATGGCAGAAACC 1264
QY 1021 ACAGATAACGCTCTTCACCTTGGATGAGATTTT 1053
Db 1265 ACAGATAACGCTCTTCACCTTGGATGAGATTTT 1297

RESULT 3

US-10-019-733-12

; Sequence 12, Application US/10019733

; Patent No. 6682917

; GENERAL INFORMATION:

; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA

; TITLE OF INVENTION: New Protein and its DNA

; FILE REFERENCE: 2619W00P

; CURRENT APPLICATION NUMBER: US/10/019,733

; CURRENT FILING DATE: 2001-12-28

; PRIOR APPLICATION NUMBER: JP 11-181131

; PRIOR FILING DATE: 1999-06-28

; PRIOR APPLICATION NUMBER: JP 11-192391

; PRIOR FILING DATE: 1999-07-06

; PRIOR APPLICATION NUMBER: JP 2000-017770

; PRIOR FILING DATE: 2000-01-21

; NUMBER OF SEQ ID NOS: 14

; SEQ ID NO 12

; LENGTH: 1053

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION:

US-10-019-733-12

Query Match 99.8%; Score 1051.4; DB 4; Length 1053;
Best Local Similarity 99.9%; Pred. No. 2.7e-310; Indels 0; Gaps 0;
Matches 1052; Conservative 0; Mismatches 1;

QY 1 ATGGGCGACCCGGAAGCGCGGCTGGATCAGGATCAGAGATCATCTTCA 60
Db 1 ATGGGCGACCCGGAAGCGCGGCTGGATCAGGATCAGAGATCATCTTCA 60
QY 61 GACACCAAGAGTGAATGAAGTCAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAAGAGTGAATGAAGTCAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
QY 121 CGGTTTGTCTATCTTCCATCCAGTACCTGATATTGGAAATGTATAACAGGCACAG 180
Db 121 CGGTTTGTCTATCTTCCATCCAGTACCTGATATTGGAAATGTATAACAGGCACAG 180
QY 181 GCTTCTCTTCTGGACACAGAGGTCGACTTATCAAGGATCTCCCTCACTGGAACAAG 240
Db 181 GCTTCTCTTCTGGACACAGAGGTCGACTTATCAAGGATCTCCCTCACTGGAACAAG 240
QY 241 CTTAAAGCAGATGAGAAGTACTTCTCTCACTCTTACATCTTGGCAGGCTGAT 300
Db 241 CTTAAAGCAGATGAGAAGTACTTCTCTCACTCTTACATCTTGGCAGGCTGAT 300
QY 301 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGCT 360
Db 301 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGCT 360
QY 361 CGCTGTTTCTATGGCTTCCAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTG 420
Db 361 CGCTGTTTCTATGGCTTCCAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA 480
QY 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTGGGCCCTTGGATGAGATAGAGATAAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTGGGCCCTTGGATGAGATAGAGATAAAA 540

QY 541 TCTACTTTTGGGAAAGAGTGGTGGCCCTTCTGCTGTAGAGAGAGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCCCTTCTGCTGTAGAGAGAGTTTCTTCTCAGGA 600
QY 601 TCTTTTGTCTATATCTGCTAAAGAGAGAGGTCTTATGCCAGGAGTCACTTTTTC 660
Db 601 TCTTTTGTCTATATCTGCTAAAGAGAGAGGTCTTATGCCAGGAGTCACTTTTTC 660
QY 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGTGCTTTGCTTGCTGATGTTCCAA 720
Db 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGTGCTTTGCTTGCTGATGTTCCAA 720
QY 721 TACTTAGTAATAAGCCTTTCAGAGAAAGGGTCAGGAGATCATTTGTGATGCTGTCAA 780
Db 721 TACTTAGTAATAAGCCTTTCAGAGAAAGGGTCAGGAGATCATTTGTGATGCTGTCAA 780
QY 781 ATTGAGCAGGAGTTTAAACAGAAAGCCCTTGGCCAGTGGCCCTCATTTGGAATGAATTGCATT 840
Db 781 ATTGAGCAGGAGTTTAAACAGAAAGCCCTTGGCCAGTGGCCCTCATTTGGAATGAATTGCATT 840
QY 841 TTGATGAACACAGTACATTGAGTTTGTAGCTGACAGATTTACTTTGTGAACTTGGATTCTCA 900
Db 841 TTGATGAACACAGTACATTGAGTTTGTAGCTGACAGATTTACTTTGTGAACTTGGATTCTCA 900
QY 901 AAGTTTTTTCAGGAGAGAAATCCTTTTGAATTTTATGAGAAACATTTCTTTAGAGGAAAA 960
Db 901 AAGTTTTTTCAGGAGAGAAATCCTTTTGAATTTTATGAGAAACATTTCTTTAGAGGAAAA 960
QY 961 ACAAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGAGTTATGCGAGAAACC 1020
Db 961 ACAAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGAGTTATGCGAGAAACC 1020
QY 1021 ACAGATAACGCTCTTCACTTGGATGAGATTTT 1053
Db 1021 ACAGATAACGCTCTTCACTTGGATGAGATTTT 1053

RESULT 4

US-10-019-733-4

; Sequence 4, Application US/10019733

; Patent No. 6682917

; GENERAL INFORMATION:

; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA

; TITLE OF INVENTION: New Protein and its DNA

; FILE REFERENCE: 2619W00P

; CURRENT APPLICATION NUMBER: US/10/019,733

; CURRENT FILING DATE: 2001-12-28

; PRIOR APPLICATION NUMBER: JP 11-181131

; PRIOR FILING DATE: 1999-06-28

; PRIOR APPLICATION NUMBER: JP 11-192391

; PRIOR FILING DATE: 1999-07-06

; PRIOR APPLICATION NUMBER: JP 2000-017770

; PRIOR FILING DATE: 2000-01-21

; NUMBER OF SEQ ID NOS: 14

; SEQ ID NO 4

; LENGTH: 1081

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION:

US-10-019-733-4

Query Match 99.8%; Score 1051.4; DB 4; Length 1081;
Best Local Similarity 99.9%; Pred. No. 2.7e-310; Indels 0; Gaps 0;
Matches 1052; Conservative 0; Mismatches 1;

QY 1 ATGGGCGACCCGGAAGCGCGGCTGGATCAGGATCAGAGATCATCTTCA 60
Db 20 ATGGGCGACCCGGAAGCGCGGCTGGATCAGGATCAGAGATCATCTTCA 79
QY 61 GACACCAAGAGTGAATGAAGTCAAGAGCCACTCTTAAAGAAAGAGTTCTCGC 120
Db 80 GACACCAAGAGTGAATGAAGTCAAGAGCCACTCTTAAAGAAAGAGTTCTCGC 139

Qy	121	CGGTTTGTCACTTTTCCAAATCCAGTACCCCTGATATTTGGAAAAATGATATAAAACAGGCACAG	180
Db	140	CGGTTTGTCACTTTTCCAAATCCAGTACCCCTGATATTTGGAAAAATGATATAAAACAGGCACAG	199
Qy	181	GCTTCCTTCTGGACAGCAGAGAAGAGGTGCACTTATCAAAGGATCTCCCTCACTCGGAACAG	240
Db	200	GCTTCCTTCTGGACAGCAGAGAAGAGGTGCACTTATCAAAGGATCTCCCTCACTCGGAACAG	259
Qy	241	CTTAAAGCAGATGAGAAGTACTTCATCTCTCACATCTTTAGCCCTTTTTTGGCAGCGAGTAT	300
Db	260	CTTAAAGCAGATGAGAAGTACTTCATCTCTCACATCTTTAGCCCTTTTTTGGCAGCGAGTAT	319
Qy	301	GGAAATGTAATGAAAAATTTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGTTCAGAGGCT	360
Db	320	GGAAATGTAATGAAAAATTTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGTTCAGAGGCT	379
Qy	361	CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAAATGTTCACTCAGAGATGTTACAGTTTG	420
Db	380	CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAAATGTTCACTCAGAGATGTTACAGTTTG	439
Qy	421	CTGATACACACTTACATCAGAGATCCCAAGAAAGGGAAATTTTATTAAATGCAATTGAA	480
Db	440	CTGATACACACTTACATCAGAGATCCCAAGAAAGGGAAATTTTATTAAATGCAATTGAA	499
Qy	481	ACCATGCCCTATGTTAAAGAAAAAGCAGATTTGGCCCTTTGCGATGGATAGCAGATAGAAAA	540
Db	500	ACCATGCCCTATGTTAAAGAAAAAGCAGATTTGGCCCTTTGCGATGGATAGCAGATAGAAAA	559
Qy	541	TCTACTTTTGGGAAAGAGTGGTGGCCCTTTGCTGCTGTGTAGAAGAGTTCCTTCTCAGGA	600
Db	560	TCTACTTTTGGGAAAGAGTGGTGGCCCTTTGCTGCTGTGTAGAAGAGTTCCTTCTCAGGA	619
Qy	601	TCITTTGCTGCTATATTTCTGGCTAAAGAGAGGCTCTTATGCGCAGACACTCACTTTTCC	660
Db	620	TCITTTGCTGCTATATTTCTGGCTAAAGAGAGGCTCTTATGCGCAGACACTCACTTTTCC	679
Qy	661	AAAGAACTCATCAGCAGAGATGAAGGACTTCACCTGTCACCTTGCTGCCCTGATGTTCCAA	720
Db	680	AATGAACTCATCAGCAGAGATGAAGGACTTCACCTGTCACCTTGCTGCCCTGATGTTCCAA	739
Qy	721	TACTTAGTAATAAGCCCTTCAGAAAGAGGTGAGGAGATCAATGTTGATGCTGTCAAA	780
Db	740	TACTTAGTAATAAGCCCTTCAGAAAGAGGTGAGGAGATCAATGTTGATGCTGTCAAA	799
Qy	781	ATTGAGCAGGAGTTTTTAAACAGAAAGCCTTGCCAGTTGGCCCTCAATGGAATGAATTCGATT	840
Db	800	ATTGAGCAGGAGTTTTTAAACAGAAAGCCTTGCCAGTTGGCCCTCAATGGAATGAATTCGATT	859
Qy	841	TTGATGAACAGTACATTTGATGTTGTAGCTGACAGATTACTTGTGGAACTTGATTTCTCA	900
Db	860	TTGATGAACAGTACATTTGATGTTGTAGCTGACAGATTACTTGTGGAACTTGATTTCTCA	919
Qy	901	AAGGTTTTTTCAGGCAGAAAAATCCCTTTTATTATTTAGTAAACATTTCTTTAGAGGAAAA	960
Db	920	AAGGTTTTTTCAGGCAGAAAAATCCCTTTTATTATTTAGTAAACATTTCTTTAGAGGAAAA	979
Qy	961	ACAAATTTCTTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGGTTATGGCAGAAACC	1020
Db	980	ACAAATTTCTTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGGTTATGGCAGAAACC	1039
Qy	1021	ACAGATAACGTCCTTCACCTTGGATGCGAGATTTT	1053
Db	1040	ACAGATAACGTCCTTCACCTTGGATGCGAGATTTT	1072

RESULT 5

RESULTS 3
US-09-962-665-9

US-09-362-665-3
; Sequence 9, Application US/09962665

; Patent No. 6537759

GENERAL INFORMATION:

APPLICANT: Stanton, Jr., Vincent P.

TITLE OF INVENTION: FOLYLPOLYGLUTAMATE SYNTHETASE GENE SEQUENCE


```
Db 817 ACTGGGCTTGCCTGATGGGACAAAGAGGCTACCTATGTTGAACGTTGTAGCCT 876
Qy 569 TTGCTGCTGAGAGGAGTTTCTCTCAGAGTCTTTTGTGCTATATCTCGGCTAAGA 628
Db 877 TTGCTGAGTGAAGGCAATTTCTTTTCGGTCTTCTTTGCGTCTGATATCTGGCTCA 936
Qy 629 AGAGAGTCTTATGCCAGGACTCACTTTTCCATGAACTCATCAGCAGAGATGAAG 688
Db 937 AACGAGGAGTGTGCTGGCTCCTCACATTTTCTAATGAACCTTATTAGCAGAGAT 996
Qy 689 TTCACTGTGACTTTGCTTGTCTGATGTTCCAACTATTAGTAAATAAGCCTTTCAG 748
Db 997 TACACTGTGATTTTGTGCTGATGTTCAACACCTGGTACACAAACCTCGAGAGA 1056
Qy 749 GGCTCAGGAGATCATTTGTTGATGCTGTCAAAATTCAGCAGGAGTTTAAACAG 808
Db 1057 GAGTAAGAGAAATAATATCAATGCTGTTCCGATAGAACAGGAGTTCTCTACT 1116
Qy 809 TGCAGTTGCGCTCATTTGGAATGAAATGCAATTTTGTATGAAACAGTACATTA 868
Db 1117 TGCCTGTGAAGCTCAITGGGATGAATTCACCTCTAATGAAGCAATACATTA 1176
Qy 869 CTGACAGATTACTTGTGGAATGGAATTCCTCAAGGTTTTCAGGAGAAATCCTTT 928
Db 1177 CAGACAGACTTATGCTGGAATGAGTTTTCAGCAAGGTTTTCAGAGTAGAACC 1236
Qy 929 ATTTATGGAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACGAGTTTC 988
Db 1237 ACTTTATGGAATATTTCTGGAAGAAAGACTAATTTCTTTGAGAGAGTAGG 1296
Qy 989 AGTATCAGCGTTTTCAGTTATGGCAGAAACCAAGATACGTTTCCACCTTGG 1048
Db 1297 AGTATCAGAGGATGGGAGTGTCAAGTCCAAACAGAGATTTCTTTTACCT 1356
Qy 1049 ATTT 1052
Db 1357 ACTT 1360
```

RESULT 6

```
US-09-963-333-9
; Sequence 9, Application US/09963333
; Patent No. 6664062
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: THYMIDINE SYNTHASE GENE SEQUENCE VARIANCES
; TITLE OF INVENTION: HAVING UTILITY IN DETERMINING THE TREATMENT
; FILE REFERENCE: OF DISEASE
; FILE REFERENCE: 11926-015002
; CURRENT APPLICATION NUMBER: US/09/963,333
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 128, 1464
; OTHER INFORMATION: n = g or a
; NAME/KEY: misc_feature
; LOCATION: 189
```

```
; OTHER INFORMATION: n = t or g
; NAME/KEY: misc_feature
; LOCATION: 524
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc_feature
; LOCATION: 1399
; OTHER INFORMATION: n = t or a
; NAME/KEY: misc_feature
; LOCATION: 1636, 1738, 2259
; OTHER INFORMATION: n = c or t
US-09-963-333-9
```

Query Match 55.0%; Score 579; DB 4; Length 2500;

Best Local Similarity 75.0%; Pred. No. 6.4e-166;

Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

```
Qy 89 ATGAGAGCCACTCCTAAGAAAGAGTTCTCGCCGCTTGTCTATCTTCCAAATCCAGTACC 148
Db 397 AGGATGAGCCCTGCTGAGAGAAACCCCGCCGCTTGTCTATCTTCCCCTCGAGTACC 456
Qy 149 CTGATATTTGAAAAATGTATAACAGGCACAGGCTTCTCTGGACAGCAGAAAGGCTCG 208
Db 457 ATGATATCTGCAGATGTATAAGAGGCAAGGCTTCTTTGGACCCGCGAGAGGTTG 516
Qy 209 ACTTATCAAAGGATCTCCCTCACTGGAACAAAGCTTTAAAGCAGATCAGAACTTCT 268
Db 517 ACCTCTNAAGGACATTGAGCACTGGGAATCCCTGAAACCCGAGGAGATATTTTATAT 576
Qy 269 CTCACTCTTAGCCCTTTTTCAGCCAGTGTGGAATTTGTAATGAAATTTTGGTGAGC 328
Db 577 CCCATGTTCTGGCTTTCTTTGCAAGCAAGCATGGCATAGTAATGAAATCTTTGGTG 636
Qy 329 GCTTTAGTTCAGGAGTGCAGGTTCCAGAGGCTCGCTGTTTCTATGCTTTCAAAATCT 388
Db 637 GATTAGCAAGAGTTGAGATTTACGAAGCCCGCTGTTCTATGGCTTCCAAATTCGCA 696
Qy 389 TCGAAGATGTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAGAT 448
Db 697 TGGAAACATACATTTCTGAATGTATAGTCTTCTTATTGACACTTACATAAAGATCC 756
Qy 449 AGAAAGGGAATTTTATTTAATGCAATTTGAAACCATGCCCCTATGTTAAGAAAAAGC 508
Db 757 AAGAAAGGGAATTTCTCTTCAATGCCATTGAAACGATGCTTGTGTCAAGAAAGAGC 816
Qy 509 ATTGGGCTTGCAGTGGATAGACATGAAATCTACTTTTGGGAAAGAGTGGTGGCT 568
Db 817 ACTGGGCTTGCCTGGATGGGAGCAAAAGGCTACCTATGTTGCAACGTTGTAGCCT 876
Qy 569 TTGCTGCTGTAGAAGGAGTTTCTTCTCAGGATCTTTTGTCTCTATATTCTGGCTAAA 628
Db 877 TTGCTGAGTGAAGGCAATTTCTTTCCGGTCTTTTCCGGTCTGATATTTCTGGCTCAA 936
Qy 629 AGAGAGTCTTATGCCAGGACTCACTTTTCCAATGAACTCATCAGCAGAGATGAAG 688
Db 937 AACGAGGAGTGTGCTGGCTCCTCATTTTCTAATGAACCTTATTAGCAGAGATGAG 996
Qy 689 TTCACTGTGACTTTGCTTGCCTGATGTTCCAACTATTAGTAAATAAGCCTTCAAGA 748
Db 997 TACACTGTGATTTTGTGCTGATGTTTCAAAACACCTGTTCAACAAACCTCGAGAG 1056
Qy 749 GGCTCAGGAGATCATTTGTTGATGCTGTCAAAATTTGAGCAGGAGTTTAAACAGAG 808
Db 1057 GAGTAAGAGAAATAATATCAATGCTGTTCCGATAGAACAGGAGTTCTCTACTGAG 1116
Qy 809 TGCCAGTTGGCCTCATTTGGAATGAAATTTGATGAAACAGTACATTTGATTTGTAG 868
Db 1117 TGCCTGTGAAGCTCATTTGGGATGAATTCACCTCTAATGAAGCAATACATTTGAT 1176
Qy 869 CTGACAGATTACTTGTGGAATCTGGAATCTCAAGGTTTTCAGGAGAGAAATCCTTT 928
Db 1177 CAGACAGACTTATGCTGGAATCTGGGTTTTCAGCAAGGTTTTCAGAGTAGAACC 1236
Qy 929 ATTTATGGAACATTTCTTTAGAGGAAACAAATTTCTTTTGGAGAAACGAGTTTC 988
```

Db 1237 ACTTTATGAGAATATTTCTACTGGAAGGAAAGACTAACTTCTTTGAGAAAGAGTAGGCG 1296
 Qy 989 AGTATCAGCGTTTTCAGTATGCGAGAAACACACAGATAACGCTTTCACCTTGGATGCGAG 1048
 Db 1297 AGTATCAGAGGATGGAGTGATGTCAAGTCCAAACAGAGAATCTTTTACCTTGGATGCTG 1356
 Qy 1049 ATTT 1052
 Db 1357 ACTT 1360

RESULT 7

US-09-962-677-9
 ; Sequence 9, Application US/09962677
 ; Patent No. 6759200
 ; GENERAL INFORMATION:
 ; APPLICANT: Stanton, Jr., Vincent P.
 ; TITLE OF INVENTION: THYMIDINE PHOSPHORYLASE GENE SEQUENCE
 ; TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING
 ; TITLE OF INVENTION: THE TREATMENT OF DISEASE
 ; FILE REFERENCE: 11926-015003
 ; CURRENT APPLICATION NUMBER: US/09/962,677
 ; CURRENT FILING DATE: 2001-09-24
 ; PRIOR APPLICATION NUMBER: 09/658,659
 ; PRIOR FILING DATE: 2000-09-08
 ; PRIOR APPLICATION NUMBER: 09/596,033
 ; PRIOR FILING DATE: 2000-06-15
 ; PRIOR APPLICATION NUMBER: 09/357,743
 ; PRIOR FILING DATE: 1999-07-20
 ; PRIOR APPLICATION NUMBER: 09/357,024
 ; PRIOR FILING DATE: 1999-07-19
 ; PRIOR APPLICATION NUMBER: 60/093,484
 ; PRIOR FILING DATE: 1998-07-20
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 9
 ; LENGTH: 2500
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 128, 1464
 ; OTHER INFORMATION: n = g or a
 ; NAME/KEY: misc_feature
 ; LOCATION: 189
 ; OTHER INFORMATION: n = t or g
 ; NAME/KEY: misc_feature
 ; LOCATION: 524
 ; OTHER INFORMATION: n = c or g
 ; NAME/KEY: misc_feature
 ; LOCATION: 1399
 ; OTHER INFORMATION: n = t or a
 ; NAME/KEY: misc_feature
 ; LOCATION: 1636, 1738, 2259
 ; OTHER INFORMATION: n = c or t
 ; US-09-962-677-9

Query Match 55.0%; Score 579; DB 4; Length 2500;
 Best Local Similarity 75.0%; Pred. No. 6.4e-166;
 Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;
 Qy 89 ATGAAGAGCCTCTTAAGAGAGTCTCGCGGTTTGTCTATCTTCCATCCAGTACC 148
 Db 397 AGGATGAGCGCTGTGAGAGAAACCCCGCGCTTGTGTCATCTTCCCATCGAGTACC 456
 Qy 149 CTGATATTTGGAATGTATAACAGGCACAGGCTTCTCTCGACAGAGAGAGTGC 208
 Db 457 ATGATATCTGGCAGATGTATAAGAGCGAGAGGCTTCTTTGACCGCGAGAGGTTG 516
 Qy 209 ACTTATCAAGAGTCTCCCTCACTGGAAACAAAGCTTAAAGCAGATGAGAAGTACTTCATCT 268
 Db 517 ACCTCTCNAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT 576

RESULT 8

US-09-949-016-2025
 ; Sequence 2025, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03

Qy 269 CTCACATCTTTAGCCTTTTTCAGCCAGTGCATGGAATGTAAATGAAAAATTTGGTGAGC 328
 Db 577 CCATGTTCTGGCTTTCTTTGACGACGATGTCATAGTAAATGAAAAATTTGGTGAGC 636
 Qy 329 GCTTTAGTCAGGAGGTCAGGTTCCAGAGGCTCGCTGTTTCTATGGCTTTCATAATCTCA 388
 Db 637 GATTTAGCCCAAGAGTTCAGATTACAGAAGCCGCTGTTCTATGGCTTCCAAATTTGCCA 696
 Qy 389 TCGAGAATGTTCACTCAGAGATGTACAGATTTGCTGTAGACACCTTACATCAGAGATCCCA 448
 Db 697 TGGAAAAACATACATTTCTGAATGTATAGTCTTCTTATGACACTTACATAAAGATCCCA 756
 Qy 449 AGAAAAAGGGAATTTTATTTAATGCAATGAAACCATGCCCCATGTTTAAAGAAAAAGCAG 508
 Db 757 AAGAAAAGGGAATTTCTCTCAATGCCATTGAAACGATGCTTGTGTCAAGAAGAGGCAG 816
 Qy 509 ATTGGGCTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 568
 Db 817 ACTGGGCTTCGCTGATTTGGGACAAAGAGGCTACCTATGTTGAAAGCTGTTGTAGGCT 876
 Qy 569 TTGCTGCTGTAGAAGGAGTTTCTTCTCAGATCTTTTGTGCTATATTTCTGGCTAAAGA 628
 Db 877 TTGCTGCACTGGAAGGCAATTTCTTTCCGGTCTTTTGGCTCGATATTTCTGGCTCAAGA 936
 Qy 629 AGAGAGGCTTTATGCCAGGACTCACTTTTTCATGAACTCATCAGCAGATGAAGAC 688
 Db 937 AACGAGGACTGATGCTGCTGCTCACAATTTCTAATGAATTTATTAGCAGAGATGAGGTT 996
 Qy 689 TTCACTGTGATTTGCTTGGCTGATGTTCCAACTTCTAGTAAATAGGCTTTCAGAGAA 748
 Db 997 TACACTGTGATTTTGTGCTGATGTTCAAAACCTGGTACACAAACCATCGAGGAGA 1056
 Qy 749 GGGTCAGGAGATCATTTGATGCTCTCAAAATTTGAGCAGAGGTTTAAACAGAGCT 808
 Db 1057 GAGTAAGAGAAATTAATTAATGCTGTTGGATAGAACAGAGTTCTCTACTGAGGCT 1116
 Qy 809 TGCCAGTTGGCCTCATTTGGAATGAATTTGATGAAACAGTACATTCAGTTTGTAG 868
 Db 1117 TGCTGTGAAGCTCATTTGGGATGAATTTGCACTTAATGAAGCAATACATTCAGTTTGTG 1176
 Qy 869 CTGACAGATTTACTGTGGAATCTTGGATCTCAAGGTTTTCAGGACGAAAAATCTTTTG 928
 Db 1177 CAGACAGACTTATGCTGGAACCTGGGTTTTCAGAGGTTTTCAGAGTAGAGAACCCATTTG 1236
 Qy 929 ATTTTATGGAAGCAATTTCTTTAGAAGGAAAAACAAATTTCTTTGAGAAACGAGTTTCAG 988
 Db 1237 ACTTTATGGAATATTTTCACTGGAAGGAAAGACTAACTTCTTTGAGAGAGATGAGCG 1296
 Qy 989 AGTATCAGCGTTTTCAGTTTATGCGAGAAACCAAGATAACGCTTCTTACCTTGGATGAG 1048
 Db 1297 AGTATCAGAGATGGGATGATGTCAAGTCCACAGAGAAATCTTTTACCTTGGATGCTG 1356
 Qy 1049 ATTT 1052
 Db 1357 ACTT 1360

; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2025
 ; LENGTH: 2479
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-09-949-016-2025

Query Match 54.9%; Score 578.4; DB 4; Length 2479;
 Best Local Similarity 75.0%; Pred. No. 9.7e-166;
 Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

QY	89	ATGAAGACCACTCTTAAGAAAGAGTTCTCGCGGTTTGTCACTCTTTCCAAATCCAGTACC	148
DB	397	AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTTGTCACTCTCCCATCGAGTACC	456
QY	149	CTGATATTTGGAAATGTATAACAGGCAAGGCTTCTTCTGGACAGCAGAGAGTGC	208
DB	457	ATGATATCTGGCAGATGTATAAGAGCGCAGAGGCTTCTTTGGACCGCGAGGAGTGG	516
QY	209	ACTTATCAAGGATCTCCCTCACTGGACAGCTTAAGCAGATGAGAGTACTTCATCT	268
DB	517	ACCTCTCCAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTATAT	576
QY	269	CTCAGATCTTAGCTTTTTCAGCCAGATGATGAATGTAAATGAAATTTGGTGAGC	328
DB	577	CCATGTTCTGGCTTTCTTTCAGCAGCAGATGCGATGATGAATGAAATTTGGTGAGC	636
QY	329	GCTTATGTCAGAGGTGCGAGTTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAATTTCA	388
DB	637	GATTTAGCAAGAGTTTCAGATTACAGAGCCCGCTGTTTCTATGGCTTCAAATTTCCA	696
QY	389	TCGAGATGTTCACTCAGAGATGACAGTTTGTCTGATAGACACTTACATCAGATCCCA	448
DB	697	TGGAATACATCACTTGAATGTATAGTCTTCTTATGACACTTACATTAAGAGATCCCA	756
QY	449	AGAAAGGGAATTTTATTAATGCAATGCAACCATGCGCTATGTTTAAAGAAAGCAG	508
DB	757	AGAAAGGGAATTTCTTCAATGCCATGAAACGATGCTTGTGTCAGAGAGGAGCAG	816
QY	509	ATTGGGCTTGGATGATGATAGAGATAAGAAATCTACTTTGGGAAAGAGTGGTGGCT	568
DB	817	ACTGGGCTTGGCTGATGATGAGGACCAAGAGGCTACTATGATGAGCGTTGTAGCT	876
QY	569	TGCTGCTGTAGAGAGTTTCTTCTCAGGATCTTTTGTGCTGATATCTTGTGCTAAAGA	628
DB	877	TTGCTGAGTGAAGGCAATTTCTTTTCGGTCTTTTGGCTCGATATCTGCTCAAGA	936
QY	629	AGAGAGTCTTATGCCAGGACTCACTTTTCCATGATCACTCAGCAGAGATGAAGAC	688
DB	937	AACGAGGACTGATGCTGCGCTCACATTTCTAATGAACCTTATGACAGAGATGAGGTT	996
QY	689	TTCACTGTCACTTTGCTGCTGATGTTTCAATCTAGTAAATAAGCCCTTCAGAGAAA	748
DB	997	TACACTGTGATTTTGTCTGCTGATGTTTCAACACCTGTTACAAACCATCGAGAGA	1056
QY	749	GGTCAAGGAGATCATTTGTGATGCTGTCAAAATTCAGCAGGAGTTTAAACAGAGCCT	808
DB	1057	GAGTAAGAGAAATATTAATCAATGCTGTTTCGGATAGAACAGGAGTTTCTCACTGAGCCT	1116
QY	809	TGCCAGTTGGCTCATTTGAATGAATTCATTTTGTATGAAACAGTACATTTGTTGTAG	868
DB	1117	TGCTGTGAAGCTCATTTGGGATGAATTCACATCTAATGAAGCAATCATTTGAGTTTGTG	1176
QY	869	CTGACAGATTTACTTTGGAACTTGGATTTCTCAAGGTTTTTTCAGGCGAGAAATCTTTTG	928
DB	1177	CAGACAGACTTATGCTGGAACTGGGTTTTAGCAAGTTTTTCAGAGTAGAGAACCCATTTG	1236
QY	929	ATTTATGAAACATTTCTTTTAGAGGAGAAACAAATTTCTTTGAGAAACGAGTTTCAG	988
DB	1237	ACTTTATGAGAAATTTCTCACTGGAAGGAAAGACTAACTTCTTTTGAAGAGAGTAGGCG	1296

QY	989	AGTATCAGCGTTTTCAGCTTATGGCAGAAACACAGATTAACCTTTCACTTGGATGCAG	1048
DB	1297	AGTATCAGAGATGGAGTGATGTCAAGTCCAACAGAGAAATCTTTTACCTTGGATGCTG	1356
QY	1049	ATTT 1052	
DB	1357	ACTT 1360	

RESULT 9
 US-09-023-655-1370
 ; Sequence 1370, Application US/09023655
 ; Patent No. 6607879
 ; GENERAL INFORMATION:
 ; APPLICANT: Cocks, Benjamin G.
 ; APPLICANT: Susan G. Stuart
 ; APPLICANT: Jeffrey J. Sellhauer
 ; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
 ; NUMBER OF SEQUENCES: 1508
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 ; STREET: 3174 PORTER DRIVE
 ; CITY: PALO ALTO
 ; STATE: CALIFORNIA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/023,655
 ; FILING DATE: HEREWITH
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Zeller, Karen J.
 ; REGISTRATION NUMBER: 37,071
 ; REFERENCE/DOCKET NUMBER: PA-0001 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (650) 855-0555
 ; TELEFAX: (650) 845-4166
 ; INFORMATION FOR SEQ ID NO: 1370:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2500 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: GENBANK
 ; CLONE: g36154
 ; US-09-023-655-1370

Query Match 54.9%; Score 578.4; DB 4; Length 2500;
 Best Local Similarity 75.0%; Pred. No. 9.7e-166;
 Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

QY	89	ATGAAGACCACTCTTAAGAAAGAGTTCTCGCGGTTTGTCACTCTTTCCAAATCCAGTACC	148
DB	397	AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTTGTCACTCTCCCATCGAGTACC	456
QY	149	CTGATATTTGGAAATGTATAACAGGCAAGGCTTCTTCTGGACAGCAGAGAGTGC	208
DB	457	ATGATATCTGGCAGATGTATAAGAGCGCAGAGGCTTCTTTTGGACCGCGAGGAGTTG	516
QY	209	ACTTATCAAGGATCTCCCTCACTGGAACAGCTTAAAGCAGATGAGAGTACTTCATCT	268

Db 517 ACCTCCAGAGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTATAT 576
Qy 269 CTCACATCTTACCTTTTTCAGCCAGTGATGGAATTTGTAATGAAATTTTGGTGGAGC 328
Db 577 CCCATGTTCTGCTTCTTTGAGCAGCATGGCATAGTAATGAAACTTTGGTGGAGC 636
Qy 329 GCTTTAGTCAGAGGTGCAAGTTCAGAGGTCGCTGTTTCTATGGCTTTCAAAATCTCA 388
Db 637 GATTTAGCCAGAAAGTTCAGATTACAGAGCCGCTGTTTCTATGGCTTTCCAAATGGCA 696
Qy 389 TCGAAGATGTCACCTCAGAGGTGATCAGTTTCTGTATAGACATACATCAGATGCCA 448
Db 697 TGAAGAACATACATCTGAAATGATAGTCTTCTTATGACACTATACATAAAGATCCCA 756
Qy 449 AGAAAGGGAATTTTATTAATGCAATTTGAACCATGCTATGTTAAAGAAAAGAGCAG 508
Db 757 AAGAAAGGGAATTTCTTCAATGCCATTTGAACGATGCTTGTCTCAGAGAGAGGCGAG 816
Qy 509 ATTGGGCTTGGATGGATGAGATGAGAAATCTACTTTTGGGGAAGAGTGGTGGCCT 568
Db 817 ACTGGGCTTGGCTGGATGGGACAAAGAGGCTACCTATGGTGAACGTTGTAGCCT 876
Qy 569 TTGCTGCTGATAGAGAGTCTTCTCAGGATCTTTTCTGCTGCTATATTTCTGGCTAAGA 828
Db 877 TTGCTGCTGATAGAGGCAATTTCTTTTCCGCTTCTTTTGGCTGATATTTCTGGCTCA 936
Qy 629 AGAGAGGCTTATGCCAGGACTCACCTTTTCAATGAACTCATCAGCAGAGATGAAGGAC 688
Db 937 AACGAGGACTGATGCTGGCCTCACATTTCTTAATGAACCTTATAGCAGATGAGGGTT 996
Qy 689 TTCACTGTGACTTTGCTTGCCTGATGTTCCAAATACCTTAGTAATAAGCCTTCAGAAAGAA 748
Db 997 TACACTGTGATTTGCTGCTGATGTTCAACACCTGGTACACAAACCATCGAGGAGA 1056
Qy 749 GGGTCAGGAGATCATTTGTGATGCTGTCATAAATTCAGCAGGAGTTTTAAACAGAGCCT 808
Db 1057 CAGTAAGAGAAATATTAATCAATGCTGTTTCGGATAGAACAGGAGTTCCTCACTGAGGCCT 1116
Qy 809 TGCCAGTTGGCCTCATTTGCAATGAAATGCAATTTTGAATGAAACAGTACATTTGAGTTGTAG 868
Db 1117 TGCCGTGAGCTCATTTGGGATGAAATGCACTCTAATGAACCAATACATTTGAGTTGTGG 1176
Qy 869 CTGACAGATTAATTTGGAACCTTGGATTTCTCAAGGTTTTTTCAGGCGAGAAATTCCTTTTG 928
Db 1177 CAGACAGACTTATGCTGGAACCTGGGTTTTAGCAAGTTTTTCAGAGTAGAGAACCCATTG 1236
Qy 929 ATTTATGAAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAACGAGTTTCAG 988
Db 1237 ACTTTATGAGAATATTTCACTGGAAGGAAAGACTAACTTTCTTTGAGAAGAGATAGGCG 1296
Qy 989 AGTATCAGCGTTTTCAGTTATGGCAGAAACCAACAGATACGCTTTCACCTTGGATGCAG 1048
Db 1297 AGTATCAGAGGATGGGAGTGAATGTCAGTCCCAACAGAGAATTTCTTTACCTTGGATGCTG 1356
Qy 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 10
US-09-949-016-145
; Sequence 145, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 145
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-145

Query Match 54.9%; Score 578.4; DB 4; Length 2500;
Best Local Similarity 75.0%; Pred. No. 9.7e-166;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

Qy 89 ATGAAGAGCCACTCTTAAGAAAGAGTTCTGCGCCGTTTGTCTATCTTTTCCAATCCAGTACC 148
Db 397 AGATGAGCGCTGCTCAGAGAAACCCCGCCGCTTGTCTATCTTTCCCATCGAGTACC 456
Qy 149 CTGATATTTGGAATAATGATAAAGAGGACAGGCTTCTCTGACAGCAGAGAGGTCG 208
Db 457 ATGATATCTGGCAGATGATAAAGAGGAGGAGGCTTCTTTTGGACCGCCGAGGAGTTG 516
Qy 209 ACTATCAAGAGATCTCCCTCACTGGAAACAGCTTTAAGCAGATGAGAGTACTTCACT 268
Db 517 ACCTCTCAAGAGCAATTCAGCACTGGGAATCCCTGNAACCCGAGGAGAGATATTTTATAT 576
Qy 269 CTCACATCTTAGCCTTTTTCAGCCAGTGAATGTAATGTAATGTAATGTAATGTAATGTAAT 328
Db 577 CCATGTTCTGGCTTTCTTTTGCAGCAAGCGATGCAATGTAATGTAATGTAATGTAATGTAAT 636
Qy 329 GCTTTAGTCAGAGGAGTTCAGGTTTCAGAGGCTGCTGTTTCTATGGCTTTCAAAATCTCA 388
Db 637 GATTTAGCCAAAGATTTTCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 696
Qy 389 TCAGAGATGTTCTCAGAGATGATGATGTTTGTCTGATAGACACTTACATCAGAGATCCCA 448
Db 697 TGAAGAAACATACATCTCGAAATGATATGTTCTTCTTATGACACTTACATAAAGATCCCA 756
Qy 449 AGAAAGGGAATTTTATTAATGCAATTTGAACCATGCTGTTTAAAGAAAAGAGCAG 508
Db 757 AAGAAAGGGAATTTCTCTTCAATGCCATTTGAACGATGCTTGTCTCAGAGAGAGGAG 816
Qy 509 ATTGGGCTTTCGATGATAGCAGATAGAAATCTACTTTTGGGGAAGAGTGGTGGCCT 568
Db 817 ACTGGGCTTTCGCTGATTTGGGCAAGAGGCTACCTATGCTGAACTGTTGTAGCT 876
Qy 569 TTGCTGCTGATAGAGGAGTCTTCTCTCAGAGATCTTTTGTCTGCTATATTTCTGGCTAAGA 828
Db 877 TTGCTGAGTGAAGGAGCAATTTCTTTTCCGCTTCTTTTGGCTGCTATATTTCTGGCTCAAGA 936
Qy 629 AGAGAGGCTTATGCCAGGACTCACCTTTTCCAAATGCAATCAGCAGAGATGAGAGGAC 688
Db 937 AACGAGACTGATGCTGGCCTCACATTTCTAATGAACCTTATAGCAGATGAGGGTT 996
Qy 689 TTCACTGTGACTTTGCTTGCCTGATGTTTCCAAATACCTTAGTAATAAGCCTTCAGAAAGAAA 748
Db 997 TACACTGTGATTTGCTTGCCTGATGTTTCAACACCTGGTACACAAACCATCGAGGAGA 1056
Qy 749 GGGTCAGGAGATCATTTGTGATGCTGTCATAAATTCAGCAGGAGTTTTAAACAGAGCCT 808
Db 1057 GAGTAAGAGAAATTAATTAATCAATGCTGTTTCGGATAGAACAGGAGTTCCTCACTGAGGCCT 1116
Qy 809 TGCAGTTGGCCTCATTTGGAATGAAATGCAATTTTGAATGAAACAGTACATTTGAGTTGTAG 868
Db 1117 TGCTGTGAAGCTCATTTGGGATGAAATGCACTCTAATGAACCAATACATTTGAGTTGTGG 1176
Qy 869 CTGACAGATTAATTTGGAACCTTGGATTTCTCAAGGTTTTTTCAGGCGAGAAATTCCTTTTG 928
Db 1177 CAGACAGACTTATGCTGGAACCTGGGTTTTAGCAAGTTTTTCAGAGTAGAGAACCCATTG 1236
Qy 929 ATTTATGAAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAACGAGTTTCAG 988
Db 1237 ACTTTATGAGAATATTTCACTGGAAGGAAAGACTAACTTTCTTTGAGAAGAGATAGGCG 1296
Qy 989 AGTATCAGCGTTTTCAGTTATGGCAGAAACCAACAGATACGCTTTCACCTTGGATGCAG 1048
Db 1297 AGTATCAGAGGATGGGAGTGAATGTCAGTCCCAACAGAGAATTTCTTTACCTTGGATGCTG 1356
Qy 1049 ATTT 1052
Db 1357 ACTT 1360

Db 1237 ACTTATGAGATATTTCACTGGAAGGAAGACTAACTTCTTTTGAGAAGAGTAGGCG 1296
 QY 989 AGTATCAGCGTTTTCAGTATGCGAGAAACCAAGATAACGTTTCCACTTGGAGCAG 1048
 Db 1297 AGTATCAGAGGATGGGAGTGATGTCAGTCCAACAGAGAATTTCTTTTACCTTGGATGCTG 1356
 QY 1049 ATTT 1052
 Db 1357 ACTT 1360

RESULT 11

US-08-905-223-125
 ; Sequence 125, Application US/08905223
 ; Patent No. 6222029
 ; GENERAL INFORMATION:
 ; APPLICANT: Edwards, Jean-Baptiste D.
 ; APPLICANT: Duclert, Aymeric
 ; APPLICANT: Lacroix, Bruno
 ; TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS
 ; NUMBER OF SEQUENCES: 503
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Knobbe, Martens, Olson & Bear
 ; STREET: 501 West Broadway
 ; CITY: San Diego
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 92101-3505
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy Disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: Win95
 ; SOFTWARE: Word
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/905,223
 ; FILING DATE:
 ; CLASSIFICATION: 536
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Israel, Ned A.
 ; REGISTRATION NUMBER: 29,655
 ; REFERENCE/DOCKET INFORMATION:
 ; TELEPHONE: (619) 235-8550
 ; TELEFAX: (619) 235-0176
 ; INFORMATION FOR SEQ ID NO: 125:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 481 base pairs
 ; TYPE: NUCLEIC ACID
 ; STRANDEDNESS: DOUBLE
 ; TOPOLOGY: LINEAR
 ; MOLECULE TYPE: CDNA
 ; ORIGINAL SOURCE:
 ; ORGANISM: Homo Sapiens
 ; TISSUE TYPE: Brain
 ; FEATURE:
 ; NAME/KEY: sig peptide
 ; LOCATION: 41..343
 ; IDENTIFICATION METHOD: Von Heijne matrix
 ; OTHER INFORMATION: score 4.4
 ; OTHER INFORMATION: seq ISHILAFPAASDG/IV
 ; US-08-905-223-125

Query Match 41.9%; Score 441; DB 3; Length 481;
 Best Local Similarity 100.0%; Pred. No. 3.2e-124;
 Matches 441; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGGCGAGCCCGAAGCGGCGGCTGGATCAGATGAGATCATCTTCA 60
 Db 41 ATGGCGAGCCCGAAGCGGCGGCTGGATCAGATGAGATCATCTTCA 100
 QY 61 GACACCAACGAAGTGAATTAAGTCAAATGAAGAGCCACTCCCTAAGAAAGAGTTCTCGC 120
 Db 101 GACACCAACGAAGTGAATTAAGTCAAATGAAGAGCCACTCCCTAAGAAAGAGTTCTCGC 160

QY 121 CGGTTTGTCTCTTCCATCCAGTACCCCTGATATTTGGAAAATGTATATAACAGGCACAG 180
 Db 161 CGGTTTGTCTCTTCCATCCAGTACCCCTGATATTTGGAAAATGTATATAACAGGCACAG 220
 QY 181 GCTTCCCTTCTCGACAGAGAGAGTCCACTTATCAAAAGGATCTCCCTCACTGGAAACAAG 240
 Db 221 GCTTCCCTTCTCGACAGAGAGAGTCCACTTATCAAAAGGATCTCCCTCACTGGAAACAAG 280
 QY 241 CTTAAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTTAGCCCTTTTTCGAGCCAGTGAT 300
 Db 281 CTTAAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTTAGCCCTTTTTCGAGCCAGTGAT 340
 QY 301 GGAATTGTAATGAAAATTTGGTGAGCGCTTTTGTAGTCAGAGGTCAGAGTTCAGAGGCT 360
 Db 341 GGAATTGTAATGAAAATTTGGTGAGCGCTTTTGTAGTCAGAGGTCAGAGTTCAGAGGCT 400
 QY 361 CGCTGTTTCTATGCTTCAAAATTTCTCATCGAAGATGTTCTCACTCAGAGATGTACAGTTTG 420
 Db 401 CGCTGTTTCTATGCTTCAAAATTTCTCATCGAAGATGTTCTCACTCAGAGATGTACAGTTTG 460
 QY 421 CTGATAGACACTTATCATCAGA 441
 Db 461 CTGATAGACACTTATCATCAGA 481

RESULT 12

US-08-307-499-1/c
 ; Sequence 1, Application US/08307499
 ; Patent No. 5651972
 ; GENERAL INFORMATION:
 ; APPLICANT: Moyer, Richard W.
 ; APPLICANT: Vi uela, Eladio
 ; APPLICANT: Gibbs, E.P.J.
 ; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
 ; NUMBER OF SEQUENCES: 60
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: David R. Saliwanchik
 ; STREET: 2421 N.W. 41st Street, Suite A-1
 ; CITY: Gainesville
 ; STATE: Florida
 ; COUNTRY: U.S.A.
 ; ZIP: 32606
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/307,499
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/908,241
 ; FILING DATE: 1-JUL-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/908,630
 ; FILING DATE: 29-JUN-1992
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/342,212
 ; FILING DATE: 21-APR-1992
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Saliwanchik, David R.
 ; REGISTRATION NUMBER: 31,794
 ; REFERENCE/DOCKET NUMBER: UP35.1.FWCC1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 904-375-8100
 ; TELEFAX: 904-372-5800
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:

```

; LENGTH: 14176 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3852..4226
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 4585..4887
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5131..5310
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5760..5912
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 6786..7130
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 10148..10513
;
; US-08-307-499-1
;
; Query Match
; Best Local Similarity 40.8%; Score 429.8; DB 1; Length 14176;
; Matches 651; Conservative 0; Mismatches 327; Indels 5; Gaps 2;
;
; QY 71 AAAGTGAATAAAGTCAAAATGAAGACCCACTCTCTAAGAAAGAGTTCTCGCCGGTTGTCA 130
; DB 10841 AAAATGATTTTATACACGAATGGAGCCCTATTCTTCAAGAGTCTGATTTCTAGGTTGGTTA 10782
;
; QY 131 TCTTTCCATCCAGTACCTCTGATTTTGGAAATGTATAAACAGGACACAGGCTCTCTCT 190
; DB 10781 TTTTCCCTATTAAATGATCATGATATCTGGAATATGTAACCAATCAGTGGCAAGTTTT 10722
;
; QY 191 GGACAGCAAGAGGTGCGACTTATCAAAAGATCTCCTCACTGGAACAAGCTTAAAGCAG 250
; DB 10721 GGACCGTGAAGAGTAGATTTATCAAAAGATTTAGATGATGGGATTAATTAACCTAAG 10662
;
; QY 251 ATGAGAAGTACTTCACTCTCTCAATCTTAGCCCTTTTTCGAGCCAGTGATGGAATGTAA 310
; DB 10661 ACGAAAATACTTTATAAAACATATACCTAGCACTTTTTCGATCTAGTGATGTTGTAA 10602
;
; QY 311 ATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGTTTCAGAGGCTCGCTGTTCT 370
; DB 10601 ATGAGAATTTAGCGAAAGATTTTATGTGATGTACAGTGTTCAGAGGCAAGATGTTCT 10542
;
; QY 371 ATGGCTTTCAAATCTCATCGAAGTTCCTCACTCAGAGATGTACAGTTGCTGATAGACA 430
; DB 10541 ATGGATTTCAAATAGCTATGGAATAATTCATTCAGAAATGTATAGTTTATTAATAGATA 10482
;
; QY 431 CTTACATCAGAGATCCCAAGAAAGGAATTTTATTAATGCAATGGAACCATGCGCT 490
; DB 10481 CATATGAGAGATATATAGAAAATGCAATTTTAAACGCTATAGAAACATGGAAT 10422
;
; QY 491 ATGTTAAGAAAAAAGCAGATTTGGCCCTTGGCATGATGATGATGATGATGATGATGATG 550
; DB 10421 GCGTAAAAAAGAAAGCTGATTTGGCCCAAGAAATGGATATC---TAGCAACAAGGTATATG 10365
;
; QY 551 GGGAAAGAGTGGTGGCTTGTGCTGTAGAGAGTTCCTTCTCAGGATCTTTTGTCTG 610
; DB 10364 GAGAAAGAGTGTAGATGATTTGCGAGCTGTGGAGGAATATTTCTTTCTGGTTCAATTTGCTG 10305
;
; QY 611 CTATATCTGGCTAAAGACAGAGGTCTTATGCGAGGACTCACTTTTCCAAATGAACTCA 670
; DB 10304 CTATATTTGGATATAAAACAGAGGATGTATGCGGATTAACATTTCTTAAGACTAA 10245
;
; QY 671 TCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 730
; DB 10244 TAAGTAGACGAAAGGTTTACATGATTTGCGTGTGATTTGCGTGTGATTTTAAACATTTATTAC 10185

```

```

QY 731 ATAAAGCTTCAGAAAGAGGTCAGGAGATCATTTGTTGATCTGCTCAAAATTCAGCAGG 790
DB 10184 ATCCACCATCTAAGGAAGTTTATAACGTCGATAATCATTTGATCGCGTTAATATAGAAAAG 10125
QY 791 AGTTTTTAACAGAAAGCCTTCGCCAGTTGCGCTCATTTGGAATGAAATTTGCAATTTGATGAAAC 850
DB 10124 AGTTTTTGACAGTTGCTATTTCCCGTGGATCTTATAGGTATGAAATTTGTTGTTAATGCTC 10065
QY 851 AGTACATTTGAGTTTGTAGCTGCAGAGATTACTTTGTGGAACCTTGGATTCTCAAGGTTTTTC 910
DB 10064 AGTATATAGAAATTCGTCGCAGATAGATTATTAACAGAGTTAGTTGTGAAAG--TCTCA 10007
QY 911 AGGAGAAAATCCTTTTGTATTTATGAAAACATTTCTTTAGAGAAAACAAATTTCT 970
DB 10006 ATGTATATATCCTTTTGTAGCTTTTATGAGTATATATCACTAGAAAGTTAAGACTAATTTT 9947
QY 971 TTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGATTATGCGAGAAACACAGATAACG 1030
DB 9946 TCGAACGACGAGTTAGTGAATATCAAAAGATGGGGGTGTTTACAAATTAAGAGAGAATA 9887
QY 1031 TCTTCACCTTGGATGCAGATTTT 1053
DB 9886 TATTTCTACGGATATAGATTTT 9864
;
; RESULT 13
; US-08-307-499-14
; Sequence 14, Application US/08307499
; Patent No. 5651972
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; TITLE OF INVENTION: Live Vaccine Vector
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: U.S.A.
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/307,499
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,241
; FILING DATE: 1-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,630
; FILING DATE: 29-JUN-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/342,212
; FILING DATE: 21-APR-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF35.1.FWCC1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14176 base pairs

```


STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: Florida

COUNTRY: U.S.A.

ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/299,268

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/901,127

FILING DATE:

APPLICATION NUMBER: US 07/908,241

FILING DATE: 1-JUL-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/908,630

FILING DATE: 29-JUN-1992

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/342,212

FILING DATE: 21-APR-1992

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Saliwanchik, David R.

REGISTRATION NUMBER: 31,794

REFERENCE/DOCKET NUMBER: UF35.1.FWCC1

TELEPHONE: 904-375-8100

TELEFAX: 904-372-5800

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 14176 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: unknown

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS

LOCATION: 3852..4226

FEATURE:

NAME/KEY: CDS

LOCATION: 4585..4887

FEATURE:

NAME/KEY: CDS

LOCATION: 5131..5310

FEATURE:

NAME/KEY: CDS

LOCATION: 5760..5912

FEATURE:

NAME/KEY: CDS

LOCATION: 6786..7130

FEATURE:

NAME/KEY: CDS

LOCATION: 10148..10513

FEATURE:

US-09-299-268-1

Query Match 40.8%; Score 429.8; DB 3; Length 14176;

Best Local Similarity 66.2%; Pred. No. 5.9e-120; Indels 5; Gaps 2;

Matches 651; Conservative 0; Mismatches 327;

QY 71 AAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGCCGGTTTGTC 130

Db 10841 AAAATGATTTTATACAGAAATGAGCCCTATTCTTCAAGAGTCTGATCTTAGTTCTGTTA 10782

QY 131 TCTTTCCAAATCCAGTACCTTGATATTGGAATATGTAATAACAGGACAGGCTTCCTTCT 190

Db 10781 TTTTCCCTATTAGTATCATGATATCTGGAATATGTAATAACCAATCAGTGGCAAGTTTTT 10722

QY 191 GGACAGCAGAAGAGGTGCGACTTATCAAGGATCTCCCTCACTGGAACAGCTTAAAGCAG 250

Db 10721 GGACCGTTGAAGAAGTAGATTATCAAAAGATTAGATGATGGGATAAATTAACATAAG 10662

QY 251 ATGAGAAGTACTTCACTCTCAATCTTACGCTTTTGGAGCCAGTCATCGAATTTGTA 310

Db 10661 ACGAATAATACTTTTATAAACAATATAGCAATTTTGGCATCTAGTGATGTTGTA 10602

QY 311 ATGAAAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCTCGCTGTTCT 370

Db 10601 ATGAGAATTTAGCGGAAGATTATGTCGATGTACAGTGTTCAGAGGACGATGTTCT 10542

QY 371 ATGCTTTTCAAAATCTCATCGAGATGTTCTACTCAGAGATGTACAGTTTGTGATAGACA 430

Db 10541 ATGGAATTTCAATAGCTATGGAATAATTCATTCAGAAATGATATGTTTATTAATAGATA 10482

QY 431 CTTACATCAAGATCCCAAGAAAGGAAATTTTATTTAATGCAATTCAGAACCATGCGCT 490

Db 10481 CATATGTAAGAGATAATATAGAAAAATGCAATTTATTTAACGCTATAGAAAACATGGAAT 10422

QY 491 ATGTTAAGAAAAAGCAGATTGGGCTTGGGATGAGATAGCAGATAGAAAAATCTACTTTTG 550

Db 10421 GCGTAAAAAGAAAGCTGATTGGGCCAGAAAAATGATATC---TAGCAACAAGGTATATG 10365

QY 551 GGGAAAGAGTGGTGGGCTTTTGTGCTGTAGAGAGTGTTCCTCTCAGGATCTTTTGTG 610

Db 10364 GAGAAAGAGTAGTAGCATTTGTCAGCTGTGGAGGAAATATCTTTTCTGTTCAATTTGCTG 10305

QY 611 CTATATTCGGCTAAGAAAGAGAGTCTTATGCGAGGCTCACCTTTTCCCAATGAACACA 670

Db 10304 CTATATTTTGGATAAAAAACGAGGATTGATGCCCGGATTAACATTTTCTAATGAACATA 10245

QY 671 TCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTGCTGATGTTCCAATCTACTAGTAA 730

Db 10244 TAAGTAGAGCAGAGGTTTACATTTGCTGTTTGGTGTGTTAATGTTTAAACATTTATTAC 10185

QY 731 ATAAGCCTTCAGAAAGAGGTTCAGGAGATCATTTGTTGATGCTGCTGCTGCTGCTGCTG 790

Db 10184 ATCCACCATCTAAGGAAGTTTATAACGTCGATTAATCAATGATGCGGTTAATATAGAAAAG 10125

QY 791 AGTTTTTAACAGAGCCTTCCAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 850

Db 10124 AGTTTTTGCAGTGTGCTTATTCGCGTGGATCTTATAGGATGAAATGTTTAAATGCTC 10065

QY 851 AGTACATTGAGTTGTAGCTGACAGATTACTTGTGGAACTTGGATTTCTCAAGGTTTTC 910

Db 10064 AGTATATAGAAATTCGTCGAGATAGATTATTAACAGAGTTAGTTGTGAAAAG--TCTCA 10007

QY 911 AGCAGAAAAATCCTTTTGTATTTTATGAAAAACATTTCTTTTGAAGAAAAACAAATTTCT 970

Db 10006 ATGTATATATCCTTTTGTAGCTTTTATGGAGTATATATCACTAGAAGGTAAGACTAATTTT 9947

QY 971 TTGAAAAACGATTTCAAGATATCAGGCTTTGTCAGTTATGCGAGAAACCAAGATAAG 1030

Db 9946 TCGAACGACGAGTTAGTGAATATCAAAAGATGCGGGGTGTTTACAAATAAAGAAAGATA 9887

QY 1031 TCTTCACTTGGATGCAATTTT 1053

Db 9886 TATTTTCTACGGATATAGATTTT 9864

RESULT 15

US-09-299-268-14

; Sequence 14, Application US/09299268

; Patent No. 6217882

; GENERAL INFORMATION:

; APPLICANT: Moyer, Richard W.

; APPLICANT: Vi uela, Eladio

; APPLICANT: Gibbs, E.P.J.

; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a

; TITLE OF INVENTION: Live Vaccine Vector

; NUMBER OF SEQUENCES: 60

; CORRESPONDENCE ADDRESS:

QY	731	ATAAGCCTTCAGAGAAAGGTCAGGAGATCATTTGATGCTGTCAAAATTTGACAGG	790
Db	3993	ATCCACCATCTAAGGAAGTTATAACGTCGATAATCATTCGCGTTAATATAGAAAAGG	4052
QY	791	AGTTTTTAACAGAGCCTTGCAGTTGGCCTCATTTGGAATGAAATTCGATTTGATGAAC	850
Db	4053	AGTTTTTGACAGTTGCTATTCGGTGGATCTTATAGGTATGAATTTGTTTAATGTCTC	4112
QY	851	AGTACATTCAGTTTGTAGCTGCAGAGATTACTTGTGGAACCTTGGATTCTCAAAAGTTTTTC	910
Db	4113	AGTATATAGAAATTCGTCGAGATAGATTATTAACAGAGTTAGTTGTGAAAAG--TCTCA	4170
QY	911	AGGCAGAAATCCTTTTGTATTTATGGAACAAATTTCTTTAGAGGAAAAACAAATTTCT	970
Db	4171	ATGTATATAATCCTTTTAGCTTTATGGAGTATATATCACTAGAAAGGTAGACTAAATTTT	4230
QY	971	TTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACACAGATAACG	1030
Db	4231	TCGAACGACGAGTTAGTGAATATCAAAAGATGGGGTGTTCACAAATAAAGNAGAGAATA	4290
QY	1031	TCTTACCTTGGATGCAGATTTT	1053
Db	4291	TATTTCTACGATATAGATTTT	4313

Search completed: June 16, 2005, 09:43:26
Job time : 224 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2005, 08:31:43 ; Search time 753 Seconds
(without alignments)
8680.672 Million cell updates/sec

Title: US-10-698-228-2
Perfect score: 1053
Sequence: 1 atggggcaccggaagcc.....tcaccttgatgcagatttt 1053

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 6054689 seqs, 3103772919 residues

Total number of hits satisfying chosen parameters: 12109378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications, NA:*

- 1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/prodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/prodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/prodata/2/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/prodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/prodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/prodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/prodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/prodata/2/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/prodata/2/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/prodata/2/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/prodata/2/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/prodata/2/pubpna/US10G_PUBCOMB.seq:*
- 20: /cgn2_6/prodata/2/pubpna/US10H_PUBCOMB.seq:*
- 21: /cgn2_6/prodata/2/pubpna/US10I_PUBCOMB.seq:*
- 22: /cgn2_6/prodata/2/pubpna/US10_NEW_PUB.seq:*
- 23: /cgn2_6/prodata/2/pubpna/US11A_PUBCOMB.seq:*
- 24: /cgn2_6/prodata/2/pubpna/US11_NEW_PUB.seq:*
- 25: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq:*
- 26: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1053	100.0	1053	18	US-10-698-228-2
2	1053	100.0	4955	17	US-10-172-118-71
3	1053	100.0	4955	18	US-10-342-887-71
4	1053	100.0	4955	18	US-10-698-228-3
5	1051.4	99.8	1053	18	US-10-698-228-12
6	1051.4	99.8	1081	18	US-10-698-228-4
7	578.4	54.9	1989	9	US-09-925-301-505
					Sequence 2, Appli
					Sequence 71, Appl
					Sequence 3, Appli
					Sequence 12, Appl
					Sequence 4, Appli
					Sequence 505, App

8	578.4	54.9	2216	15	US-10-084-817-342	Sequence 342, App
9	578.4	54.9	2500	9	US-09-954-456-724	Sequence 724, App
10	578.4	54.9	2500	9	US-09-954-456-1169	Sequence 1169, App
11	578.4	54.9	2500	9	US-09-954-456-1827	Sequence 1827, App
12	578.4	54.9	2500	18	US-10-641-643-1370	Sequence 1370, Ap
13	578.4	54.9	2500	20	US-10-733-878-458	Sequence 458, App
14	578.4	54.9	2500	21	US-10-843-641A-3751	Sequence 3751, Ap
15	578.4	54.9	2500	21	US-10-843-641A-4136	Sequence 4136, Ap
16	578.4	54.9	2500	21	US-10-843-641A-4854	Sequence 4854, Ap
17	578.4	54.2	2113	21	US-10-764-420-1636	Sequence 1636, Ap
18	561	53.3	1328	18	US-10-403-571-75	Sequence 75, Appl
19	545.4	51.8	977	17	US-10-264-237-790	Sequence 790, App
20	447.8	42.5	186854	21	US-10-872-156-34	Sequence 34, Appl
c	442	42.0	1146	9	US-09-822-830A-174	Sequence 174, App
21	432.2	41.0	1218	15	US-10-128-714-7245	Sequence 7245, Ap
22	432.2	41.0	1218	15	US-10-128-714-7245	Sequence 34052, A
23	403.2	38.3	963	17	US-10-369-493-34052	Sequence 2245, Ap
24	397	37.7	1146	15	US-10-128-714-2245	Sequence 46252, A
25	392.6	37.3	1173	17	US-10-369-493-46252	Sequence 6537, A
26	392	37.2	1242	16	US-10-032-585-6537	Sequence 27809, A
27	386.6	36.7	1292	17	US-10-369-493-27809	Sequence 126843,
28	380.6	36.1	1248	18	US-10-424-599-126843	Sequence 45857, A
29	380.4	36.1	1200	17	US-10-369-493-45857	Sequence 40471, A
30	371.2	35.3	1361	18	US-10-424-599-40471	Sequence 73024, A
31	365.8	34.7	1586	18	US-10-424-599-73024	Sequence 2145, Ap
32	364.8	34.6	1206	17	US-10-320-797-2145	Sequence 43, Appl
33	361.6	34.3	957	20	US-10-653-047-43	Sequence 1245, Ap
34	349	33.1	1314	15	US-10-128-714-1245	Sequence 6245, Ap
35	349	33.1	1450	15	US-10-128-714-6245	Sequence 245, App
36	349	33.1	3314	15	US-10-128-714-245	Sequence 5245, Ap
37	349	33.1	3450	15	US-10-128-714-5245	Sequence 12381, A
38	328.2	31.2	560	9	US-09-864-761-12381	Sequence 31361, A
39	324	30.8	1169	18	US-10-425-114-31361	Sequence 31182, A
40	324	30.8	2992	20	US-10-425-115-31182	Sequence 15462, A
41	318.6	30.3	1569	20	US-10-425-115-15462	Sequence 1378, Ap
42	314.4	29.9	1306	20	US-10-739-930-1378	Sequence 13723, A
43	312.8	29.7	1255	19	US-10-767-701-13723	Sequence 63321, A
44	309.6	29.4	1791	20	US-10-425-115-63321	Sequence 11109, A
c	306.2	29.1	1725	19	US-10-437-963-11109	

ALIGNMENTS

RESULT 1

US-10-698-228-2
; Sequence 2, Application US/10698228
; Publication No. US20040072253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619W00P
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-2

Query Match 100.0%; Score 1053; DB 18; Length 1053;
Best Local Similarity 100.0%; Pred. No. 7.9e-295;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	601	TC	TTTTGCTGCTATATTTCTGGCTAAAGAGAGAGGTTCTTATGCCAGGACTCACTTTTTC	660
Db	845	TC	TTTTGCTGCTATATTTCTGGCTAAAGAGAGAGGTTCTTATGCCAGGACTCACTTTTTC	904
Qy	661	AAT	GAACTCATCAGCAGAGATGAAGGACTTCACGTGCACTTTGGTTCCTGATGTTCCAA	720
Db	905	AAT	GAACTCATCAGCAGAGATGAAGGACTTCACGTGACTTTGCTTTGCCGTGATGTTCCAA	964
Qy	721	TACT	TAGTAAATAAGCCTTCAGAAAGAAAGGTCAGGGAGATCAATGTTGATGCTGTCAAA	780
Db	965	TACT	TAGTAAATAAGCCTTCAGAAAGAAAGGTCAGGGAGATCAATGTTGATGCTGTCAAA	1024
Qy	781	ATT	GAGCAGGAGTTTTTAAACAGAAAGCCTTGCCAGTTGGCCTCATTTGGAATGAATTCGATT	840
Db	1025	ATT	GAGCAGGAGTTTTTAAACAGAAAGCCTTGCCAGTTGGCCTCATTTGGAATGAATTCGATT	1084
Qy	841	TTG	ATGAAACAGTACATTTGATGTTGTAGCTGACAGATTAATCTGTGGAACTTGGAATTC	900
Db	1085	TTG	ATGAAACAGTACATTTGATGTTGTAGCTGACAGATTAATCTGTGGAACTTGGAATTC	1144
Qy	901	AAG	TTTTTTTCAGCGAGAAATCCCTTTTGATTTTATGGAACAACATTTCTTTAGAGGAAA	960
Db	1145	AAG	TTTTTTTCAGCGAGAAATCCCTTTTGATTTTATGGAACAACATTTCTTTAGAGGAAA	1204
Qy	961	ACAAA	TTCTTTTCAGAGAAACGAGTTTCAGAGTATCAGCGTTTTCGAGTTATGGCAGAGAAC	1020
Db	1205	ACAAA	TTCTTTTCAGAGAAACGAGTTTCAGAGTATCAGCGTTTTCGAGTTATGGCAGAGAAC	1264
Qy	1021	ACAG	ATAACGCTTTCACCTTGGATGACGATTTT	1053
Db	1265	ACAG	ATAACGCTTTCACCTTGGATGACGATTTT	1297

RESULT 3

```

US-10-342-887-71
; Sequence 71, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Barnards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 71
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-71

```

	Query Match	100.0%;	Score 1053;	DB 18;	Length 4955;
	Best Local Similarity	100.0%;	Pred. No. 28-294;		
	Matches 1053;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATCGGCGACCCGAAAGCGCGGAGCGCGCGCTCGATCAGAGATGAGATCATCTTCA	60		
Db	245	ATCGGCGACCCGAAAGCGCGGAGCGCGCGCTCGATCAGAGATGAGATCATCTTCA	304		
Qy	61	GACACCAAGAAAGTGAATTAAGTCAATGAAGAGCCACTCCTCAAGAAAGAGTCTTCGC	120		

305	Db	 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGGTCTCGC	364
121	Qy	CGGTTTGTCTATCTTTCCAAATCCAGTACCCCTGATATTTGGAAAAATGTAATAAACAGGCACAG	180
365	Db	 CGGTTTGTCTATCTTTCCAAATCCAGTACCCCTGATATTTGGAAAAATGTAATAAACAGGCACAG	424
181	Qy	GCTTCCCTTTGGGACAGCAGAAGAGGTGCACTTATCAAAGGATCTCCCTCATCTGGAAACAAG	240
425	Db	 GCTTCCCTTTGGGACAGCAGAAGAGGTGCACTTATCAAAGGATCTCCCTCATCTGGAAACAAG	484
241	Qy	CTTAAAGCAGATGAGAAGTACTTCACTCTCAACATCTTAGCCCTTTTTCGAGCCAGTGAT	300
485	Db	 CTTAAAGCAGATGAGAAGTACTTCACTCTCAACATCTTAGCCCTTTTTCGAGCCAGTGAT	544
301	Qy	GGAAATGTAAATCAAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAAGTTCAGAGGCT	360
545	Db	 GGAAATGTAAATCAAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAAGTTCAGAGGCT	604
361	Qy	CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGATGTCTCACTCAGAGATGTACAGTTTG	420
605	Db	 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGATGTCTCACTCAGAGATGTACAGTTTG	664
421	Qy	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTTATTAAATGCAATTTGAA	480
665	Db	 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTTATTAAATGCAATTTGAA	724
481	Qy	ACCATGCCCTATGTTTAAGAAAAAAGCAGATTTGGGCCCTTGCAGTGTGATAGCAGATAGAAAA	540
725	Db	 ACCATGCCCTATGTTTAAGAAAAAAGCAGATTTGGGCCCTTGCAGTGTGATAGCAGATAGAAAA	784
541	Qy	TCCTACTTTTGGGAAAGAGTGGTGGCTTGTCTGTGTAGAAAGAGTTTTCTCTCAGGA	600
785	Db	 TCCTACTTTTGGGAAAGAGTGGTGGCTTGTCTGTGTAGAAAGAGTTTTCTCTCAGGA	844
601	Qy	TCCTTTGTCTCTATATTTCTGGCTTAAAGAGAGAGGTCCTTATGCGAGGACTCCTTTTTTC	660
845	Db	 TCCTTTGTCTCTATATTTCTGGCTTAAAGAGAGAGGTCCTTATGCGAGGACTCCTTTTTTC	904
661	Qy	AATGMACTCATCAGCAGAGATGAAGNCTTCACTGTGACTTTCGCTGCCCTGATGTTCCAA	720
905	Db	 AATGMACTCATCAGCAGAGATGAAGNCTTCACTGTGACTTTCGCTGCCCTGATGTTCCAA	964
721	Qy	TACTTACTAAATTAAGCCCTTCAGAAGAAAGGCTCAGGAGATCAATTTGTGTGCTGTCAAA	780
965	Db	 TACTTACTAAATTAAGCCCTTCAGAAGAAAGGCTCAGGAGATCAATTTGTGTGCTGTCAAA	1024
781	Qy	ATTGAGCAGAGATTTTTTAAACAGAACCTTCGCAGTTGGCCCTCATTTGGAATGAATTCGATT	840
1025	Db	 ATTGAGCAGAGATTTTTTAAACAGAACCTTCGCAGTTGGCCCTCATTTGGAATGAATTCGATT	1084
841	Qy	TTGATGAAACAGTACATTGATGTTTGTAGCTGCACAGATTACTTGTGGAATTTGGAATTCCTCA	900
1085	Db	 TTGATGAAACAGTACATTGATGTTTGTAGCTGCACAGATTACTTGTGGAATTTGGAATTCCTCA	1144
901	Qy	AAGTTTTTTCAGGCAGAAAAATCCCTTTTGTATTTTATGGAACAATTTCTTTTAGAAGGAAAA	960
1145	Db	 AAGTTTTTTCAGGCAGAAAAATCCCTTTTGTATTTTATGGAACAATTTCTTTTAGAAGGAAAA	1204
961	Qy	ACAAAATTTCTTTCAGAAAAACAGTTTTCAGAGTATCAGCGTTTTTCAGTATATGGCGAAACC	1020
1205	Db	 ACAAAATTTCTTTCAGAAAAACAGTTTTCAGAGTATCAGCGTTTTTCAGTATATGGCGAAACC	1264
1021	Qy	ACAGATAACGCTCTTCACCTTTGGATGCAATTTT	1053
1265	Db	 ACAGATAACGCTCTTCACCTTTGGATGCAATTTT	1297

RESULT 4
US-10-698-228-3
; Sequence 3, Application US/10698228
; Publication No. US20040072253A1
; GENERAL INFORMATION:

; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
 ; TITLE OF INVENTION: New Protein and its DNA
 ; FILE REFERENCE: 2619WOOP
 ; CURRENT APPLICATION NUMBER: US/10/698,228
 ; CURRENT FILING DATE: 2003-10-30
 ; PRIOR APPLICATION NUMBER: US/10/019,733
 ; PRIOR FILING DATE: 2001-12-28
 ; PRIOR APPLICATION NUMBER: JP 11-181131
 ; PRIOR FILING DATE: 1999-06-28
 ; PRIOR APPLICATION NUMBER: JP 11-192391
 ; PRIOR FILING DATE: 1999-07-06
 ; PRIOR APPLICATION NUMBER: JP 2000-017770
 ; PRIOR FILING DATE: 2000-01-21
 ; NUMBER OF SEQ ID NOS: 14
 ; SEQ ID NO 3
 ; LENGTH: 4955
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION:
 US-10-698-228-3

Query Match 100.0%; Score 1053; DB 18; Length 4955;
 Best Local Similarity 100.0%; Pred. No. 2e-294;
 Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGGGCGACCCGGAAGCGCGAGCGCGCGGCTGGATCAGATCAGAGATCATCTTCA	60
DB	245	ATGGGCGACCCGGAAGCGCGAGCGCGCGGCTGGATCAGATCAGAGATCATCTTCA	304
QY	61	GACACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTCTCGC	120
DB	305	GACACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTCTCGC	364
QY	121	CGTTTGTTCATCTTCCCAATCCAGTACCTCTCATATTTGGAAATGATATAACAGGCACAG	180
DB	365	CGTTTGTTCATCTTCCCAATCCAGTACCTCTCATATTTGGAAATGATATAACAGGCACAG	424
QY	181	GCTTCCTCTTGGACAGACGAGAGTGCATCTCAATGAAGATCTCCCTCACTGGAACAG	240
DB	425	GCTTCCTCTTGGACAGACGAGAGTGCATCTCAATGAAGATCTCCCTCACTGGAACAG	484
QY	241	CTTAAAGCAGATCAGAGAGTACTTCTCATCTCATCTTACCTTTTGGCAGCAGTGAT	300
DB	485	CTTAAAGCAGATCAGAGAGTACTTCTCATCTCATCTTACCTTTTGGCAGCAGTGAT	544
QY	301	GGAAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT	360
DB	545	GGAAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT	604
QY	361	CGCTGTTCTATGGCTTTCAATCTCATCGAAGTTCATCTCAGAGATGTACAGTTTG	420
DB	605	CGCTGTTCTATGGCTTTCAATCTCATCGAAGTTCATCTCAGAGATGTACAGTTTG	664
QY	421	CTCATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGAATTTGAA	480
DB	665	CTCATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGAATTTGAA	724
QY	481	ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCTTTGGATGATGATGATGATGATAA	540
DB	725	ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCTTTGGATGATGATGATGATGATAA	784
QY	541	TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGATGAGAGAGTTTCTTCTCAGGA	600
DB	785	TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGATGAGAGAGTTTCTTCTCAGGA	844
QY	601	TCTTTTGTCTATATCTGGCTAAAGACAGAGGTCTTATGCCAGCACTCACTTTTTC	660
DB	845	TCTTTTGTCTATATCTGGCTAAAGACAGAGGTCTTATGCCAGCACTCACTTTTTC	904
QY	661	AATGAATCATCAGCAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA	720
DB	905	AATGAATCATCAGCAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA	964

QY	721	TACTTAGTAAATAAGCCTTCAGAAAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAA	780
DB	965	TACTTAGTAAATAAGCCTTCAGAAAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAA	1024
QY	781	ATTGAGCAGAGAGTTTAAACAGAGCCCTTCCAGTTCGCTCATTTGGAATGAATTCATT	840
DB	1025	ATTGAGCAGAGAGTTTAAACAGAGCCCTTCCAGTTCGCTCATTTGGAATGAATTCATT	1084
QY	841	TTGATGAACACAGTACATTGAGTTTGTAGCTGACAGATTACTTTGGAACTTGGATTTCTCA	900
DB	1085	TTGATGAACACAGTACATTGAGTTTGTAGCTGACAGATTACTTTGGAACTTGGATTTCTCA	1144
QY	901	AAGGTTTTTTCAGGCAGAAAAATCCTTTTGAATTTTATGGAAAAATTTCTTTTAAAGGAAAA	960
DB	1145	AAGGTTTTTTCAGGCAGAAAAATCCTTTTGAATTTTATGGAAAAATTTCTTTTAAAGGAAAA	1204
QY	961	ACAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGAGTATGCGCAGAAACC	1020
DB	1205	ACAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGAGTATGCGCAGAAACC	1264
QY	1021	ACAGATAACGCTCTTCACTTTGGATGACAGATTTT	1053
DB	1265	ACAGATAACGCTCTTCACTTTGGATGACAGATTTT	1297

RESULT 5

US-10-698-228-12
 ; Sequence 12, Application US/10698228
 ; Publication No. US20040072253A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
 ; TITLE OF INVENTION: New Protein and its DNA
 ; FILE REFERENCE: 2619WOOP
 ; CURRENT APPLICATION NUMBER: US/10/698,228
 ; CURRENT FILING DATE: 2003-10-30
 ; PRIOR APPLICATION NUMBER: US/10/019,733
 ; PRIOR FILING DATE: 2001-12-28
 ; PRIOR APPLICATION NUMBER: JP 11-181131
 ; PRIOR FILING DATE: 1999-06-28
 ; PRIOR APPLICATION NUMBER: JP 11-192391
 ; PRIOR FILING DATE: 1999-07-06
 ; PRIOR APPLICATION NUMBER: JP 2000-017770
 ; PRIOR FILING DATE: 2000-01-21
 ; NUMBER OF SEQ ID NOS: 14
 ; SEQ ID NO 12
 ; LENGTH: 1053
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION:
 US-10-698-228-12

Query Match 99.8%; Score 1051.4; DB 18; Length 1053;
 Best Local Similarity 99.9%; Pred. No. 2.3e-294;
 Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	ATGGCGCACCCGGAAGCGCGGAGCGCGGCTGGATCAGATCAGAGATCATCTTCA	60
DB	1	ATGGCGCACCCGGAAGCGCGGAGCGCGGCTGGATCAGATCAGAGATCATCTTCA	60
QY	61	GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC	120
DB	61	GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC	120
QY	121	CGTTTGTCTATCTTTCCAAATCCAGTACCTCATATTTGGAAAAATGTATAAACAGGCACAG	180
DB	121	CGTTTGTCTATCTTTCCAAATCCAGTACCTCATATTTGGAAAAATGTATAAACAGGCACAG	180
QY	181	GCTTCTCTTTCGACAGCAGAGAGTGCATTTATCAAAAGGATCTCCCTCATCTGGAACAAG	240
DB	181	GCTTCTCTTTCGACAGCAGAGAGTGCATTTATCAAAAGGATCTCCCTCATCTGGAACAAG	240

QY 241 CTTAAAGCAGATGAGAGTACTTCTCATCTCTCATCTTACGCTTTTTCAGCCAGTGAT 300
DB 241 CTTAAAGCAGATGAGAGTACTTCTCATCTCTCATCTTACGCTTTTTCAGCCAGTGAT 300
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGAGTGACAGGTTCCAGAGCT 360
DB 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGAGTGACAGGTTCCAGAGCT 360
QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCAGAAATGTTTCACTCAGAGATGTACAGTTG 420
DB 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCAGAAATGTTTCACTCAGAGATGTACAGTTG 420
QY 421 CTGATAGACATTTACATCAGAGATGCCAGAAAGGAAATTTTATTAATGCAATGAA 480
DB 421 CTGATAGACATTTACATCAGAGATGCCAGAAAGGAAATTTTATTAATGCAATGAA 480
QY 481 ACCATGCCCTATCTTAAGAAAAGCAGATTTGGCCCTTGGATGATAGCAGATAGAAA 540
DB 481 ACCATGCCCTATCTTAAGAAAAGCAGATTTGGCCCTTGGATGATAGCAGATAGAAA 540
QY 541 TCTACTTTTGGGAAAAGAGTGGCTTGGCTTGTCTGTAGAAAGGAGTTTCTTCTCAGGA 600
DB 541 TCTACTTTTGGGAAAAGAGTGGCTTGGCTTGTCTGTAGAAAGGAGTTTCTTCTCAGGA 600
QY 601 TCTTTTGTCTATATTTCTGGCTTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
DB 601 TCTTTTGTCTATATTTCTGGCTTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
QY 661 AATGAATCATCAGCAGAGATGAAGACTTCACTGTGACATTTGCTGTAGAGAGTTTCTTCTCAG 720
DB 661 AATGAATCATCAGCAGAGATGAAGACTTCACTGTGACATTTGCTGTAGAGAGTTTCTTCTCAG 720
QY 721 TACTTAGTAAATAAGCTTTCAGAAAGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAAA 780
DB 721 TACTTAGTAAATAAGCTTTCAGAAAGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAAA 780
QY 781 ATTGAGCAGAGTTTTTAAACAGAAAGCTTGCAGTTGGCTCATTTGGAATGAATTCAT 840
DB 781 ATTGAGCAGAGTTTTTAAACAGAAAGCTTGCAGTTGGCTCATTTGGAATGAATTCAT 840
QY 841 TTGATGAACAGTACATTTAGTTTGTAGCTGACAGATTAATTTGGAAATCTTGAATTTCA 900
DB 841 TTGATGAACAGTACATTTAGTTTGTAGCTGACAGATTAATTTGGAAATCTTGAATTTCA 900
QY 901 AAGGTTTTTCAGCAGAAATCTTTTGAATTTTATGGAACATTTCTTTAGAGGAAA 960
DB 901 AAGGTTTTTCAGCAGAAATCTTTTGAATTTTATGGAACATTTCTTTAGAGGAAA 960
QY 961 ACAAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020
DB 961 ACAAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020
QY 1021 ACAGATAAGCTTTTCACTTACCTTGGATGAGATTTT 1053
DB 1021 ACAGATAAGCTTTTCACTTACCTTGGATGAGATTTT 1053

RESULT 6

US-10-698-228-4

; Sequence 4, Application US/10698228
; Publication No. US20040072253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619W00P
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06

; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 4
; LENGTH: 1081
; TYPE: DNA
; ORGANISM: Homo sapiens
; OTHER INFORMATION:
US-10-698-228-4

Query Match 99.8%; Score 1051.4; DB 18; Length 1081;

Best Local Similarity 99.9%; Pred. No. 2.4e-294;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGGCAGCCGGAAGCGCGGAGCGGCTTGGATCAGGATGAGAGATCATCTTCA 60
DB 20 ATGGGGCAGCCGGAAGCGCGGAGCGGCTTGGATCAGGATGAGAGATCATCTTCA 79
QY 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTGC 120
DB 80 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTGC 139
QY 121 CGGTTTGTCTATTTTCAATCCAGTACCTCATATTTTGGAAAAATGTATAAACAGGCACAG 180
DB 140 CGGTTTGTCTATTTTCAATCCAGTACCTCATATTTTGGAAAAATGTATAAACAGGCACAG 199
QY 181 GCTTCCCTTCTGGACAGCAGAGAGTGCATTTATCAAGAGATCTCCCTCCTCGGAACAG 240
DB 200 GCTTCCCTTCTGGACAGCAGAGAGTGCATTTATCAAGAGATCTCCCTCCTCGGAACAG 259
QY 241 CTTAAAGCAGATGAGAGTACTTCTCATCTCTCACATCTTTCAGCCTTTTTCAGCCAGTGAT 300
DB 260 CTTAAAGCAGATGAGAGTACTTCTCATCTCTCACATCTTTCAGCCTTTTTCAGCCAGTGAT 319
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGAGTGACAGGTTCCAGAGCT 360
DB 320 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGAGTGACAGGTTCCAGAGCT 379
QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAAGATGTTTCACTCAGAGATGTACAGTTG 420
DB 380 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAAGATGTTTCACTCAGAGATGTACAGTTG 439
QY 421 CTGATAGACATTTACATCAGAGATCCCAAGAAAAAGGAAATTTTATTAATGCAATGAA 480
DB 440 CTGATAGACATTTACATCAGAGATCCCAAGAAAAAGGAAATTTTATTAATGCAATGAA 499
QY 481 ACCATGCCCTATTTGTAAGAAAAAGCAGATTTGGGCTTTCGATGGATAGCAGATAGAAA 540
DB 500 ACCATGCCCTATTTGTAAGAAAAAGCAGATTTGGGCTTTCGATGGATAGCAGATAGAAA 559
QY 541 TCTACTTTTGGGAAAAGAGTGGGCTTTGCTGCTGTAGAGAGTTTCTTCTCAGGA 600
DB 560 TCTACTTTTGGGAAAAGAGTGGGCTTTGCTGCTGTAGAGAGTTTCTTCTCAGGA 619
QY 601 TCTTTTGTCTATTTCTGCTTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTTCC 660
DB 620 TCTTTTGTCTATTTCTGCTTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTTCC 679
QY 661 AATGAATCATCAGCAGAGATGAAGACTTCACTGTGATGCTTTCGCTGATGTTCCAA 720
DB 680 AATGAATCATCAGCAGAGATGAAGACTTCACTGTGATGCTTTCGCTGATGTTCCAA 739
QY 721 TACTTAGTAAATAAGCTTTCAGAAAGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAA 780
DB 740 TACTTAGTAAATAAGCTTTCAGAAAGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAA 799
QY 781 ATTGAGCAGAGTTTTTAAACAGAAAGCTTGCAGTTGGCTCATTTGGAATGAATTCAT 840
DB 800 ATTGAGCAGAGTTTTTAAACAGAAAGCTTGCAGTTGGCTCATTTGGAATGAATTCAT 859
QY 841 TTGATGAACAGTACATTTGATGTTGTAGCTGACAGATTAATTTGTGGAATCTTGGATTTCA 900

Db 860 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGAACTTGGATTCTCA 919
 QY 901 AAGGTTTTTTCAGGACAGAAATCCTTTTGTATGTAAGAAACATTTCTTTAGAAAGAAAA 960
 Db 920 AAGGTTTTTTCAGGACAGAAATCCTTTTGTATGTAAGAAACATTTCTTTAGAAAGAAAA 979
 QY 961 ACAAATTTCTTGGAGAACAGTTTCAGAGTATCAGCGTTTTCAGATTATGGCAGAAACC 1020
 Db 980 ACAAATTTCTTGGAGAACAGTTTCAGAGTATCAGCGTTTTCAGATTATGGCAGAAACC 1039
 QY 1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053
 Db 1040 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1072

RESULT 7

US-09-925-301-505
 ; Sequence 505, Application US/09925301
 ; Patent No. US20020052308A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
 ; FILE REFERENCE: PA106
 ; CURRENT APPLICATION NUMBER: US/09/925,301
 ; CURRENT FILING DATE: 2001-08-10
 ; PRIOR APPLICATION NUMBER: PCT/US00/05882
 ; PRIOR FILING DATE: 2000-03-08
 ; PRIOR APPLICATION NUMBER: 60/124,270
 ; PRIOR FILING DATE: 1999-03-12
 ; NUMBER OF SEQ ID NOS: 1694
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 505
 ; LENGTH: 1989
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1917)
 ; OTHER INFORMATION: n equals a,t,g, or c
 US-09-925-301-505

Query Match 54.9%; Score 578.4; DB 9; Length 1989;
 Best Local Similarity 75.0%; Pred. No. 9.3e-157;
 Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

QY 89 ATGAAGACCACTCTTAAGAAAGAGTTCTCGCGGTTTGTCTATCTTTCCAAATCCAGTACC 148
 Db 280 AGGATGAGCGCTGTCTGAGAGAAACCCCGCGCTTTGTCTATCTTTCCCATCGAGTACC 339
 QY 149 CTGATATTTGGAAAAATGTATAACAGGACAGAGCTTCTCTTGGACAGAGAGAGTGC 208
 Db 340 ATGATATCTGGCAGATGTATAGAGGACAGAGCTTCTTTTGGACCGCGGAGAGTGC 399
 QY 209 ACTTATCAAGAGTCTCCCTCACTGAGAACAGCTTAAAGCAGATGAGAGTACTTCACT 268
 Db 400 ACCTCTCCAAGACATTTAGACCTGGGAATCCCTGAAACCCGAGGAGAGATATTTATAT 459
 QY 269 CTCACATCTAGCCTTTTGGAGCCAGTGTGAATGTAAATGAAATTTGGTGGAGC 328
 Db 460 CCCATGTTCTGGCTTTCTTGGACGAGCGATGATGATGATGATGATGATGATGATGATG 519
 QY 329 GCTTTAGTCAGGAGTGCAGGTTCCAGAGGCTCGCTTTCTATGCTTCCAAATTTCTCA 388
 Db 520 GATTTAGCCAGAGAGTTGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 579
 QY 389 TCGAGATATGTTCACTCAGAGATGTACAGTTTGTGTATGATGATGATGATGATGATGATG 448
 Db 580 TCGAAACATACATTTCTGAAATGTATAGTCTTCTTATGACACTTACATAAAGATCCCA 639
 QY 449 AGAAAGGGAATTTTATTTATGCAATTTGAAACCATGCGCTTATGTTAGAAAAAGACAG 508
 Db 640 AAGAAAGGGAATTTCTCTTCAATGCAATTTGAAACCATGCGCTTGTGTCAAGAGAGGACG 699

QY 509 ATTGGCCCTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 568
 Db 700 ACTGGCCCTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 759
 QY 569 TTCTCTCTGTAGAGAGAGATTTTCTTCTCAGGATCTTTTGTCTGTATATTTCTGGCTAAAGA 628
 Db 760 TTCTCTCTGTAGAGAGAGATTTTCTTCTCAGGATCTTTTGTCTGTATATTTCTGGCTCAAGA 819
 QY 629 AGAGAGGTCTTATGCGCAGGACTCATTCTTCAATGAACTCATCAGCAGAGATGAAAGAC 688
 Db 820 AACGAGGACTGATGCGCTGGCTTCACTTTTCTAATGAACTTATTAGCAGAGATGAGGTTT 879
 QY 689 TTCACTGTGACTTTGCTTGGCTGATGTTTCAATGAACTTATTAGCAGAGATGAGGTTT 748
 Db 880 TACACTGTGATTTTGTCTGCTGATGTTTCAACACCTGGTACACAAACCATCGGAGGAGA 939
 QY 749 GGGTCAGGAGAGATCATTTGTTGATGCTGCTCAAAATTTAGCAGGAGTCTTTTAAACAGAGCT 808
 Db 940 GAGTAAAGAAATTAATTTATCAATGCTGCTGGATAGAACAGGAGTTCTCTACTGAGGCT 999
 QY 809 TGCAGTTGGCCTCATTTGGAAATGCAATTTTGTATGAAACAGTACATCAGTGTGTTGAG 868
 Db 1000 TGCCTGTGAAGCTCATTTGGGATGAAATTTGCACTCTAATGAAGCAATACATTGAGTTTGTG 1059
 QY 869 CTGACAGATTTACTTGTGGAATCTTCAAGGTTTTCAGGAGGAGAAATCCTTTTGTG 928
 Db 1060 CAGACAGACTTATGCTGGAATCTGCTGTTTGTAGCAAGGTTTTCAGAGTATAGAACCCATTTG 1119
 QY 929 ATTTTATGAAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACGAGTTTTCAG 988
 Db 1120 ACTTTATGGAATATTTCTGAGGAAAGCACTAATCTTTTGTAGAGAGATGAGGCG 1179
 QY 989 AGTATCAGCGCTTTTGCAGTTTATGCGAGAAACCCAGATAAGCTTCTTCACTTGGATGAG 1048
 Db 1180 AGTATCAGAGGATGCGAGTGTCTGATGCTCAAGTCCACAGAGATTTCTTTTACCTTGGATGCTG 1239
 QY 1049 ATTT 1052
 Db 1240 ACTT 1243

RESULT 8

US-10-084-817-342
 ; Sequence 342, Application US/10084817
 ; Publication No. US20030119009A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Susan Stuart
 ; APPLICANT: Jed G. Nuchtern
 ; APPLICANT: Sharon E. Plon
 ; APPLICANT: Jason M. Shohet
 ; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
 ; FILE REFERENCE: PA-0046 US
 ; CURRENT APPLICATION NUMBER: US/10/084,817
 ; CURRENT FILING DATE: 2002-02-25
 ; PRIOR APPLICATION NUMBER: 60/270,784
 ; PRIOR FILING DATE: 2001-02-23
 ; NUMBER OF SEQ ID NOS: 365
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 342
 ; LENGTH: 2216
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incyte ID No. US20030119009A1 202239.1
 US-10-084-817-342

Query Match 54.9%; Score 578.4; DB 15; Length 2216;
 Best Local Similarity 75.0%; Pred. No. 9.9e-157;
 Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;
 QY 89 ATGAAGACCACTCTTAAGAAAGAGTTCTCGCGGTTTGTCTATCTTTCCAAATCCAGTACC 148

Db 284 AGGATGAGCGCTGCTGAGAGAAAAACCCCGCGCTTTGTGTCATCTTCCCATCGAGTACC 343
 QY 149 CTGATATTTGGAAATGTATAACACAGGCAAGGCTTCTCTGGAAGAGAGAGTGC 208
 Db 344 ATGATATCTGGCAGATGTATAAGAGCGCAGAGGCTTCTTTTGGACCGCGAGAGGTGG 403
 QY 209 ACTTATCAAGAGTCTCCCTCACTGGAACAAGCTTAAGCAGATGAGAAGTACTTCACT 268
 Db 404 ACCTCTCCAAGGACATCTCAGCACTGGGAATCCCTGAAACCGAGGAGAGATATTTATAT 463
 QY 269 CTCACATCTTAGCGCTTTTTCAGCCAGTGTGAATTTGAAATGAAATTTGGTGAGC 328
 Db 464 CCATGTTCTGGCTTTCTTTCAGCAGCAAGCGATGGCATAGTAATGAATCTTGGTGAGC 523
 QY 329 GCTTTAGTCAGAGGTCAGAGTTCCAGAGGCTCGCTGTTTCTATGCTTTCAAAATCTCA 388
 Db 524 GATTTAGCCAAGAGTTCAGATACAGAGCCGCTGTTTCTATGGCTTCCAAATGGCA 583
 QY 389 TCAGAGATGTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGATCCCA 448
 Db 584 TGGAAACATACATCTGAAATGTATAGTCTTCTTATGACACTTACATAAAGATCCCA 643
 QY 449 AGAAAAGGGAATTTTATTAATTAATGCAATTCGAAACCATGCCCTATGTTAAGAAAAAGCAG 508
 Db 644 AAGAAAGGGAATTTCTTCAATGCCATTCGAAAGATGCCCTTGTGTCAGAGAGGCGAG 703
 QY 509 ATTGGGCTTTCGATGATAGCAGATAGAAAAATCTACTTTGGGGAAGAGTGGTGCCT 568
 Db 704 ACTGGGCTTTCGCTGGAATGGGACAAAGAGGCTACCTATGTTGAACGTGTTAGCCT 763
 QY 569 TTGCTGCTGTAGAGAGTTTCTTCTCAGGATCTTTTGTGCTATATTTCTGCTAAGA 628
 Db 764 TTGCTGCTGAGTGAAGGCAATTTCTTTTCCGGTCTTTTGGTGCATATCTGCTCAAGA 823
 QY 629 AGAGAGTCTTATGCCAGGACTCACATTTTCCAAATGACATCATCAGCAGAGATCAAGGAC 688
 Db 824 AAGCAGGACTGATGCTGCGCTCCACATTTTCTAATGAACCTTATGACAGAGATGAGGGTT 883
 QY 689 TTCACGTGCACTTTGCTGCTGATGTTCCAACTTCTAGTAAATAAGCCTTTCAGAGAAA 748
 Db 884 TACACTGTGATTTTGTCTGCTGATGTTTCAACCTGTTACAAACCTTCGAGGAGA 943
 QY 749 GGTTCAGGAGATCAATTTGTGATGCTGTCAAAAATTGAGCAGGAGTTTAAACAGAGCCT 808
 Db 944 GAGTAAGAGAAATAATTAATCAATGCTGTTCCGATAGAACAGGAGTTTCTCACTGAGCCT 1003
 QY 809 TGGCAGTTGGCCTCATTTGGAATGAAATTTGATGATAAAGTACATTTGATGAGTGTG 868
 Db 1004 TGCCTGTGAAGCTCATTTGGGATGAATGCACTCTAATGAAGCAATACATTTGAGTTGTG 1063
 QY 869 CTGACAGATTTACTTTGGAACTTGGATTTCTCAAGGTTTTCAGGCGAGAAAAATCCTTTTG 928
 Db 1064 CAGACAGACTTATGCTGGAACCTGGGTTTATGCAAGGTTTTCAGAGTAGAGAACCCATTTG 1123
 QY 929 ATTTATGAAACATTTCTTTAGAGGAAAAACAAATTTCTTTGAGAAACGAGTTTCAG 988
 Db 1124 ACTTTATGAGATATTTCACTGGAAGGAAAGACTTAATCTTTGAGAGAGAGTAGGCG 1183
 QY 989 AGTATCAGGCTTTTGGAGTTATGGCAGAAACCAAGATACGTCCTTCACTTGGATGCGAG 1048
 Db 1184 AGTATCAGAGGATGGGAGTGATGTCAGTCCCAACAGAGAAATCTTTTACCTTGGATGCTG 1243
 QY 1049 ATTT 1052
 Db 1244 ACTT 1247

RESULT 9

US-09-954-456-724
 ; Sequence 724, Application US/09954456
 ; Patent No. US20020115057A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Young, Paul

; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Can
 ; TITLE OF INVENTION: Sets
 ; FILE REFERENCE: 689290-76
 ; CURRENT APPLICATION NUMBER: US/09/954,456
 ; CURRENT FILING DATE: 2001-09-18
 ; PRIOR APPLICATION NUMBER: US/60/233,617
 ; PRIOR FILING DATE: 2000-09-18
 ; PRIOR APPLICATION NUMBER: US/60/234,052
 ; PRIOR FILING DATE: 2000-09-20
 ; PRIOR APPLICATION NUMBER: US/60/234,923
 ; PRIOR FILING DATE: 2000-09-25
 ; PRIOR APPLICATION NUMBER: US/60/235,134
 ; PRIOR FILING DATE: 2000-09-25
 ; PRIOR APPLICATION NUMBER: US/60/235,637
 ; PRIOR FILING DATE: 2000-09-26
 ; PRIOR APPLICATION NUMBER: US/60/235,638
 ; PRIOR FILING DATE: 2000-09-26
 ; PRIOR APPLICATION NUMBER: US/60/235,711
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: US/60/235,720
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: US/60/235,840
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: US/60/235,863
 ; PRIOR FILING DATE: 2000-09-27
 ; NUMBER OF SEQ ID NOS: 2276
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 724
 ; LENGTH: 2500
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-954-456-724

Query Match 54.9%; Score 578.4; DB 9; Length 2500;
 Best Local Similarity 75.0%; Pred. No. 1.1e-156;
 Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;
 QY 89 ATGAAGAGCCACTCTCTAAGAAAGAGTCTCCCGGTTTGTGTCATCTTTTCCAATCCAGTACC 148
 Db 397 AGATGAGCGCTGCTGAGAGAAACCCCGCGCTTGTGTCATCTTTCCCATCGAGTACC 456
 QY 149 CTGATATTTGGAAATGTATAACACAGGCAAGGCTTCTCTGGAAGAGAGGTCG 208
 Db 457 ATGATATCTGGCAGATGTATAAGAGGCGAGAGGCTTCTTTTGGACCGCGAGAGGTG 516
 QY 209 ACTTATCAAGAGTCTCCCTCACTGGAACAAGCTTAAGCAGATGAGAAGTACTTCACT 268
 Db 517 ACCTCTCCAAGGACATTCAGCACTGGGAATCCCTGNAACCCGAGGAGAGATATTTTATAT 576
 QY 269 CTCACATCTTAGCCTTTTTCAGCCAGTGTGAATTTGTAATGAAATTTGGTGGAGC 328
 Db 577 CCATGTTCTGGCTTTCTTTCAGCAAGCGATGCGATAGTAATGAAATTTGGTGGAGC 636
 QY 329 GCTTTAGTCAGAGGTCAGGTTCCAGAGGCTCGCTGTTCTTCTATGGCTTTCAAAATTTCA 388
 Db 637 GATTTAGCCAAGAGTTTCAGATTACAGAAAGCCGCTGTTTCTATGGCTTCCAAATTTGCCA 696
 QY 389 TCAGAGATGTTCACTCAGAGATGTACAGTTTGTGATAGACACTTACATCAGAGATCCCA 448
 Db 697 TGGAAACATACATCTCTGAAATGTATAGTCTTCTTATGACACTTACATAAAGATCCCA 756
 QY 449 AGAAAAGGGAATTTTATTAATTAATGCAATTCGAAACCATGCCCTATGTTAAGAAAAAGCAG 508
 Db 757 AAGAAAGGGAATTTCTCTTCAATGCCATTTGAACGATGCCCTTGTGTCAGAGAGAGGCGAG 816
 QY 509 ATTGGGCTTTCGATGATAGCAGATAGAAAAATCTACTTTTGGGGAAGAGTGGTGGCT 568
 Db 817 ACTGGGCTTTCGCTGATTTGGGACAAAGAGGCTACCTATGTTGGAACGTGTTGTAGCCT 876
 QY 569 TTGCTGCTGTAGAGGAGTTTCTTCTCAGAGATCTTTTGTGCTATATTTCTGGCTAAGA 628
 Db 877 TTGCTGCAAGTGAAGGCAATTTCTTTTCCGGTCTTTTTCGCTCGATATTTCTGGCTCAAGA 936

QY	629	AGAGAGGCTTTATGCCAGGACTCATTCTTTTCCAAATGAACTCATCAGCAGAGATGAAGGAC	688
Db	937	AACCGAGGACTGATGCTTGGCCTCACAATTTTCTAAATGAACTTATTAGCAGAGATGAGGGTT	996
QY	689	TTCACTGTGACTTTGCTTGGCCTGATGTTCCAAATACCTTAGTAAATAAGCCTTCAGAAGAAA	748
Db	997	TACACTGTGATTTTGTCTTGGCCTGATGTTCAACACCTGGTACACAAACCATCGGAGGAGA	1056
QY	749	GGGTCAAGGAGATCAATTTGTATGCTGTCAAAATTTGAGCAGGAGCTTTTAAACAGAAGCCT	808
Db	1057	GAGTAAGAGAAATANTATCAATGCTGTTCGGATAGAACAGGAGTTCCTCACTGAGGCT	1116
QY	809	TGCAGTTGGCCTCAATTTGGAATGAATTTGATGAAACAGTACATTGAGTTTGTAG	868
Db	1117	TGCTGTGAAGCTCAATTTGGGATGAATTTGCACCTCTAATGAAGCAATACATTGAGTTTGTGG	1176
QY	869	CTGACAGATTACTTTGGGAACTTTGGATTCTCAAGGTTTTTTCAGGCAGAAAATCCTTTTG	928
Db	1177	CAGACAGACTTATGCTGGAACTGGGTTTTTAGCAAGGTTTTTTCAGAGTAGAAGCAATTTG	1236
QY	929	ATTTTATGCGAAACATTTCTTTTGAAGGAGAAAACAAATTTCTTTTGAGAAACGAGTTTCAG	988
Db	1237	ACTTTATGAGAAATATTTCACTGGAAGGAAAGACTAACTTCTTTTGAGAAAGAGATGAGCG	1296
QY	989	AGTATCAGCGTTTTCGACGTTATGGCAGAAACCAACAGATAACGTCCTTCACTTGGATGCAG	1048
Db	1297	AGTATCAGAGGATGGAGTGATGTCAAGTCCAAACAGAGAATTTCTTTTACCTTGGATGCTG	1356
QY	1049	ATTT 1052	
Db	1357	ACTT 1360	

RESULT 10
 US-09-954-456-1169
 ; Sequence 1169, Application US/09954456
 ; Patent No. US20020115057A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Young, Paul
 ; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Can
 ; TITLE OF INVENTION: Sets
 ; FILE REFERENCE: 689290-76
 ; CURRENT APPLICATION NUMBER: US/09/954,456
 ; CURRENT FILING DATE: 2001-09-18
 ; PRIORITY APPLICATION NUMBER: US/60/233,617
 ; PRIORITY FILING DATE: 2000-09-18
 ; PRIORITY APPLICATION NUMBER: US/60/234,052
 ; PRIORITY FILING DATE: 2000-09-20
 ; PRIORITY APPLICATION NUMBER: US/60/234,923
 ; PRIORITY FILING DATE: 2000-09-25
 ; PRIORITY APPLICATION NUMBER: US/60/235,134
 ; PRIORITY FILING DATE: 2000-09-25
 ; PRIORITY APPLICATION NUMBER: US/60/235,637
 ; PRIORITY FILING DATE: 2000-09-26
 ; PRIORITY APPLICATION NUMBER: US/60/235,638
 ; PRIORITY FILING DATE: 2000-09-26
 ; PRIORITY APPLICATION NUMBER: US/60/235,711
 ; PRIORITY FILING DATE: 2000-09-27
 ; PRIORITY APPLICATION NUMBER: US/60/235,720
 ; PRIORITY FILING DATE: 2000-09-27
 ; PRIORITY APPLICATION NUMBER: US/60/235,840
 ; PRIORITY FILING DATE: 2000-09-27
 ; PRIORITY APPLICATION NUMBER: US/60/235,863
 ; PRIORITY FILING DATE: 2000-09-27
 ; NUMBER OF SEQ ID NOS: 2276
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 1169
 ; LENGTH: 2500
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-954-456-1169

Best Local Similarity		75.0%	Pred. No. 1.1e-156;						
Matches		723;	Conservative	0;	Mismatches	241;	Indels	0;	Gaps
QY	89	ATGAAGAGCCACCTCCTAAGAAAGAGTTCTCGCGGTTTGTCACTCTTTCCAAATCCAGTACC	148						
DB	397	AGGATGAGCCGCTGCTGAGAGAAAACCCCGCGCTTTGTTCATCTTCCCATCGAGTACC	456						
QY	149	CTGATATTTGGAAAAATGTATATAACAGGCA CAGGCTTCTCTGGACAGCAGAGAAGGTCG	208						
DB	457	ATGATATCTGGCAGATGTATAAAGAGCAGAGGCTTCTTTTGGACCGCGAGGAGGTTG	516						
QY	209	ACTTATCAAAAGGATCCTCCTCAGCTGGAAACAAGCTTAAAGCAGATGAGAAGTACTTCATCT	268						
DB	517	ACCTCTCAAGGACATTCAGCACTGGGATCCCTGAAAACCCGAGGAGAGATATTTTATAT	576						
QY	269	CTCACATCTTTAGCCTTTTTTTCGACCAGTGATGGAATTTGTAATGAAAAATTTGGTGGAGC	328						
DB	577	CCCATGTTCTGCTTTCTTTTGACAGCAAGCATGGCATAGTAATGAAACTTGGTGGAGC	636						
QY	329	GCTTTAGTCAGGAGTGCGAGGTTCCAGAGGCTCGCTGTTCTTATGCGTTTCAAAATCTCA	388						
DB	637	GAITTAGCCAAAGATTTTCAGATTCAGAAAGCCGCGTTTCTATGCGTTTCCAAATTTGCCA	696						
QY	389	TCGAGAATGTTCTACTCAGAGATGTACAGTTTGTGTGTAGACACTTACATCAGAGATCCCCA	448						
DB	697	TGAAAAACATACATTTGAAATGTATAGTCTTTTATGTACACTTACATAAAAAGATCCCCA	756						
QY	449	AGAAAAAGGAAATTTTATTTAATTCGAATTCGAACCATGCCCTATGTTAAGAAAAAGCAG	508						
DB	757	AAGAAAAGGAAATTTCTCTCTCAATGCCATTGAAACGATGCCCTGTGTCAAGAAAGAGCAG	816						
QY	509	ATTGGGCTTTCGATGATGATGATGAGAAAATCTACTTTTGGGGAAGAGTGGTGGCCT	568						
DB	817	ACTGGGCTTTTCGCTGGATTTGGGGACAAAGAGGCTACCTATGTGTGAACGTTGTGTAGCCT	876						
QY	569	TTGCTGCTGTAGAAGAGATTTTCTCTCTCAGATCTTTTGTGCTGTATATTTCTGGCTAAAGA	628						
DB	877	TTGCTGCAGTGGAAAGCAATTTTCTTTTCCGGTTCTTTTGGCTCGATATTTCTGGCTCAAGA	936						
QY	629	AGAGAGGCTTTATGCCCAGGACTCACTTTTTTCCAAATGAACCTCATCAGCAGAGATGAAGGAC	688						
DB	937	AACGAGGACTGATGCCCTGGCTCACAATTTTCTAATGAACCTTATTAGCAGAGATGAGGGTT	996						
QY	689	TTCACTGTGACTTTGCTTGGCTGATGTTTCCAAATCTTAGTAATAGCCCTTCAGAGAAA	748						
DB	997	TACACTGTGATTTTGTCTGTGCTGATGTTCAACACCTGGTACAAACCAACCATCGGAGGAGA	1056						
QY	749	GGGTTCAGGAGATCATTTGTTGTGCTCTCAAAATTTGAGCAGGAGTTTAAAAACAAGACCT	808						
DB	1057	GAGTAAGAGAAATATTTATCAATGCTGTTTCGGATAGAACAGAGTTCCTCACTGAGGCT	1116						
QY	809	TGCCAGTTGGCCTCATTTGGAATGAATTCGATTTTGAATGAAAACAGTACATTTGAGTTGTAG	868						
DB	1117	TGCTGTGAAGCTCATTTGGGATGAATTCGACTCTAATGAAGCAATACATTTGAGTTGTGG	1176						
QY	869	CTGACAGATTACTTGTGGAATTTGGATTTCTCAAGGTTTTTTCAGGCAGAAAAATCCTTTTG	928						
DB	1177	CAGACAGACTTATGCTGGAACTGGTTTTTATAGCAAGGTTTTTTCAGAGTAGAAGACCATTTG	1236						
QY	929	ATTTTATGAAAAACATTTCTTTAGAAGAAAAACAAATTTCTTTGAAAAACGAGTTTTCAG	988						
DB	1237	ACTTTATGGAGATATTTTCACTGGNAGGAAGACTAATCTTTTGGAGAGAGTAGGCG	1296						
QY	989	AGTATCAGCGTTTTTGCAGTTATGGCAGAAAACCAAGATAACGCTTTCACCTTGGATCGAG	1048						
DB	1297	AGTATCAGAGGATGGGAGTGATGTCAAGTCCAAACAGAGAAATCTTTTACCTTGGATGCTG	1356						
QY	1049	ATTT	1052						
DB	1357	ACTT	1360						

US-09-954-456-1827
; Sequence 1827, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canceled Sequences
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR FILING DATE: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 1827
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-1827

Query Match 54.9%; Score 578.4; DB 9; Length 2500;
Best Local Similarity 75.0%; Pred. No. 1.1e-156;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

QY 89 ATGAAGAGCCACTCTTAAGAAAGAGTTCTCGCGGTTGTTCATCTTTCCAAATCCAGTACC 148
DB 397 AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTTGTCACTCCCTCCATCGAGTACC 456

QY 149 CTGATATTTGGAATATTAACAGCAGCAGGCTTCTCTGAGCAGCAGAGAGTGC 208
DB 457 ATGATATCTGGCAGATGTATAAGAGGCGAGAGGCTTCTTTGGACCGCGAGAGGTTG 516

QY 209 ACTTATCAAGGATCTCCCTCACTGGAACAGCTTAAAGCAGATGAGAGTACTTCACT 268
DB 517 ACCTCTCAAGGACATTCAGCATCTGGGATCCCTGAAACCGAGGAGAGATATTTATAT 576

QY 269 CTCACATCTTAGCTTTTTCAGCCAGTGAATTTGAAATGAAATTTGGTGGAGC 328
DB 577 CCATGTTCTGGCTTTCTTTGAGCAGCAGCATGGCATGATGAAATGAAATTTGGTGGAGC 636

QY 329 GCTTTAGTCAGGAGGTCAGGTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAATCTCA 388
DB 637 GATTTAGCCAAGAAGTTCAGATTACAGAGCCGCTGTTTCTATGGCTTCCAAATGGCCA 696

QY 389 TCGAGATGTTCACTCAGAGATGATGATGTTGCTGATAGACATTCATCAGATCCCA 448
DB 697 TGAAGAACATACATCTGAAATGATAGTCTTCTTATGACATTCATCAATAAAGATCCCA 756

QY 449 AGAAGAGGGAATTTTATTAATGCAATGAAACACCTGCTTATGAAAGAAAGAGCAG 508
DB 757 AGAAGAGGGAATTTCTTCAATGCCATGAAACGATGCTTGTGTGCAAGAGAGCAG 816

QY 509 ATTGGGCTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 568
DB 817 ACTGGGCTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 876

QY 569 TTGCTGCTGTAGAGAGGAGTTTCTTCTCAGGATCTTTTCTGCTCTATATTTCTGCTAAAGA 628
DB 877 TTGCTGCTGTAGAGGAGGATTTTCTTTTCCGGTCTTTTTCGGTCGATATTTCTGCTCAAGA 936

QY 629 AGAGAGGTTCTATGCCAGGAGTCTCACTTTTCCAAATGAACTCATCAGCAGAGATGAAGGAC 688
DB 937 AACGAGGACTGATGCTGCTGCTCACTTTTCTAATGAACTTATTAAGCAGAGATGAGGTT 996

QY 689 TTCACTGTGACTTTTGTGCTTGTGCTGATGTTCCCAATCTTATGTAATTAAGCCTTCAGAGAAA 748
DB 997 TACACTGTGATTTTGTGCTTGTGCTGATGTTCAACACCTGTTACACAAACCATCGAGGAGA 1056

QY 749 GGGTCAGGAGATCATTTGTTGATGCTGTCAAAATTTGAGCAGAGTTTTAAACAGAGCCT 808
DB 1057 GAGTAAGAGAAATAATTAATCAATGCTGTTCCGATAGAAACAGAGTTCCTCACTGAGGCT 1116

QY 809 TGCCAGTTGGCTCATTTGGAATGATTTGATTTGATCAAAACAGTACATTTGATTTGTAG 868
DB 1117 TGCCTGTGAAGCTCATTTGGGATGAAATTCACCTCTAATGAAGCAATACATTTGATTTGTG 1176

QY 869 CTGACAGATTACTTGTGGAACCTTGGAATCTCAAGGTTTTCAGGCAGAGAAATCCCTTTTG 928
DB 1177 CAGACAGCTTATGCTGGAATGCTGTTTTCAGAGGTTTTCAGAGTAGAGAACCCATTTG 1236

QY 929 ATTTTATGGAACCAATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACGAGTTTCAG 988
DB 1237 ACTTTATGGAGATATTTCTCACTGGAAGGAAAGACTAACTCTTTTTCAGAGAGAGTAGGCG 1296

QY 989 AGTATCAGGTTTTCAGAGTTATGAGGAGAAACCAAGATGATGATGATGATGATGATGATGAT 1048
DB 1297 AGTATCAGAGGATGGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1356

QY 1049 ATTT 1052
DB 1357 ACTT 1360

RESULT 12
US-10-641-643-1370
; Sequence 1370, Application US/10641643
; Publication No. US20040077003A1
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; Susan G. Stuart
; Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
; GENE EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/641,643
; FILING DATE: 14-Aug-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:

```

;
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1370:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 2500 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; IMMEDIATE SOURCE:
;   LIBRARY: GENBANK
;   CLONE: g36154
; SEQUENCE DESCRIPTION: SEQ ID NO: 1370 :
US-10-641-643-1370

```

```

Query Match      54.9%; Score 578.4; DB 18; Length 2500;
Best Local Similarity 75.0%; Pred. No. 1.1e-156;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

QY 89 ATGAAGAGCCATCTCTTAAGAAAGAGTTCTCGCGGTTTGTCTATCTTCCATCCAGTACC 148
Db 397 AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTGTCTATCTTCCCATCGAGTACC 456
QY 149 CTGATATTTGGAATGATAACAGGCACAGGCTTCTTCTGGACAGAGAGGTCG 208
Db 457 ATGATATCTGCGAGATGATAAGAGGCGAGAGGCTTCTTCTGGACCGCGAGGTTG 516
QY 209 ACTTATCAAGAGATCTCCCTCACTGGAACAGCTTAAAGCAGAGATGAGAGTACTCT 268
Db 517 ACCTCTCAAGAGATCTCAGCACTTGGCAATCCCTGAAACCCGAGGAGATATTTATAT 576
QY 269 CTACATCTTAGCGCTTTTTCAGCCAGTGAATGTAATGAAATTTGGTGGAGC 328
Db 577 CCATGTTCTGGCTTCTTTCGACCAAGCGATGCAATGAAATTTGGTGGAGC 636
QY 329 GCTTTAGTCAGGAGTGCAGGTTCCAGAGGCTCGCTGTTTCTATGGCTTTTCAAAATTTCTCA 388
Db 637 GATTTAGCAAGAGATGATCAATGAGAGCGGCTGTTTCTATGGCTTTCCAAATTTGCCA 696
QY 389 TCAGAGATGTTCACTCAGAGATGTACAGTTTGTGTGATAGACACTTACATCAGAGATCCCA 448
Db 697 TGAAGAAACATACATCTCTGAAATGTATAGTCTTCTTATGACACTTACATAAAGATCCCA 756
QY 449 AGAAAGGGAATTTTATTTAATGCAATTTGAACCATGCCCCCTATGTTTAAGAAAGAGCAG 508
Db 757 AAGAAAGGGAATTTCTCTCAATGCCATGAAACGATGCGCTTGTGTCAAGAGAGGCGAG 816

```

```

Query Match      54.9%; Score 578.4; DB 20; Length 2500;
Best Local Similarity 75.0%; Pred. No. 1.1e-156;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

QY 89 ATCAAGAGCCATCTCTTAAGAAAGAGTTCTCGCGGTTTGTCTATCTTCCATCCAGTACC 148
Db 397 AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTGTCTATCTTCCCATCGAGTACC 456
QY 149 CTGATATTTGGAATGATAAAGCAGGCACAGGCTTCTTCTGGACAGAGAGGTCG 208
Db 457 ATGATATCTGCGAGATGATAAGAGGCGAGAGGCTTCTTCTGGACCGCGAGGTTG 516
QY 209 ACTTATCAAGAGATCTCCCTCACTGGAACAGCTTAAAGCAGAGATGAGAGTACTCT 268
Db 517 ACCTCTCAAGAGATCTCAGCACTTGGCAATCCCTGAAACCCGAGGAGATATTTATAT 576
QY 269 CTACATCTTAGCGCTTTTTCAGCCAGTGAATGTAATGAAATTTGGTGGAGC 328
Db 577 CCATGTTCTGGCTTCTTTCGACCAAGCGATGCAATGAAATTTGGTGGAGC 636
QY 329 GCTTTAGTCAGGAGTGCAGGTTCCAGAGGCTCGCTGTTTCTATGGCTTTTCAAAATTTCTCA 388
Db 637 GATTTAGCAAGAGATGATCAATGAGAGCGGCTGTTTCTATGGCTTTCCAAATTTGCCA 696
QY 389 TCAGAGATGTTCACTCAGAGATGTACAGTTTGTGTGATAGACACTTACATCAGAGATCCCA 448
Db 697 TGAAGAAACATACATCTCTGAAATGTATAGTCTTCTTATGACACTTACATAAAGATCCCA 756
QY 449 AGAAAGGGAATTTTATTTAATGCAATTTGAACCATGCCCCCTATGTTTAAGAAAGAGCAG 508
Db 757 AAGAAAGGGAATTTCTCTCAATGCCATGAAACGATGCGCTTGTGTCAAGAGAGGCGAG 816

```

Qy	509	ATTGGGCTTGGCATGGATACGATAGAAAATCTACTTTTGGGGAAGAAGTGTGGCCT	568
Db	817	ACTGGGCTTTGGCTGGATTTGGGACAAAGAGGCTACTATGTGTGAACGTGTCTAGCCT	876
Qy	569	TTGCTGCTGTAGAAGGAGTTTCTTCTCAGGATCTTTTGTCTGTATATTCTGGCTTAAGA	628
Db	877	TTGCTGAGTGGAAAGGCAATTTCTTTCCGGTCTCTTTTGGCTGCATATTTCTGGCTCAAGA	936
Qy	629	AGAGAGTCTTATGCCAGGACTCACTTTTTCCAATGAACCTCATCAGCAGAGATGAAGGAC	688
Db	937	AACGAGGACTGATGCCCTGGCCTCACATTTTCTAATGAACCTATTAGCAGAGATCAGGGTT	996
Qy	689	TTCACTGTGACTTTGCTTTGGCTGTATGTTTCCAAATCTTAGTAATAATAGCCTTCAGAGAAA	748
Db	997	TACACTGTGATTTTGTCTTGCTGTATGTTTCAAACAACCTGGTACAACAACCATCGGAGGAGA	1056
Qy	749	GGGTCCGGGAGATCATTTGTTGATGCTGTCAAAATTTGACGAGGAGTTTTTTAACAGAACCT	808
Db	1057	GAGTAAGAGAAATTAATTATCAATGCTGTTCGGATAGAACAGGAGTTCCTCACTGAGCCT	1116
Qy	809	TGCCAGTTGGCCTCATTTGGAAATGAATTTGCATTTTGTATGAAACACGATACATTCAGTTTCTAG	868
Db	1117	TGCCTGTGAAGCTCATTTGGGATGAATTTGCATCTCTAATGAAGCAATACATTCAGTTTCTGG	1176
Qy	869	CTGACAGATTACTTGTGGAACCTTGGATTCTCAAAAGTTTTTTCAGGCGAGAAAATCCTTTTG	928
Db	1177	CAGACAGACTTATGCTCGAACTGGTTTTAGCAAGTTTTTTCAGAGTAGAGAACCCATTTG	1236
Qy	929	ATTTTATGGAACAATTTCTTTAGAAGGAAAAAATAATTTCTTTTGAAAAAGAGTTTCAG	988
Db	1237	ACTTTATGGAGAAATTTTCACTGGAAGGAAGAAGACTTAACCTCTTTTGAAGAGAGTAGGCG	1296
Qy	989	AGTATCAGCGTTTTTGCAAGTTATGGCAGAAAACCAAGATAACGCTCTTCACCTTTGGATCGAG	1048
Db	1297	AGTATCAGAGGATGGGAGTGTATGCTAGTCCACAGAGAAATCTTTTACCTTTGGATGCTG	1356
Qy	1049	ATTT 1052	
Db	1357	ACTT 1360	

RESULT 14

```

RESUBMIT 14
US-10-843-641A-3751
; Sequence 3751, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals, Inc.
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; TITLE OF INVENTION: Signature Gene Sets
; FILE REFERENCE: 689290-189
; CURRENT APPLICATION NUMBER: US/10/843,641A
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/873,367
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US/09/954,531
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,832
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/964,824
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/09/967,768
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/09/968,007
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,708
; PRIOR FILING DATE: 2001-10-03
; Remaining Prior Application data removed - See File Wrapper or PALM.

```

Db 1297 AGTATCAGAGGATGGGAGTGATGTCAAGTCCAAACAGAGAATTTCTTTTACCTTTGGATGCTG 1356
 QY 1049 ATTT 1052
 Db 1357 ACTT 1360

RESULT 15
 US-10-843-641A-4196
 ; Sequence 4196, Application US/10843641A
 ; Publication No. US2005064454A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Avalon Pharmaceuticals, Inc.
 ; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
 ; FILE OF INVENTION: Signature Gene Sets
 ; FILE REFERENCE: 689290-189
 ; CURRENT APPLICATION NUMBER: US/10/843,641A
 ; CURRENT FILING DATE: 2004-05-12
 ; PRIOR APPLICATION NUMBER: US/09/873,367
 ; PRIOR FILING DATE: 2001-06-05
 ; PRIOR APPLICATION NUMBER: US/09/954,531
 ; PRIOR FILING DATE: 2001-09-18
 ; PRIOR APPLICATION NUMBER: US/09/954,456
 ; PRIOR FILING DATE: 2001-09-25
 ; PRIOR APPLICATION NUMBER: US/09/962,436
 ; PRIOR FILING DATE: 2001-09-25
 ; PRIOR APPLICATION NUMBER: US/09/962,832
 ; PRIOR FILING DATE: 2001-09-25
 ; PRIOR APPLICATION NUMBER: US/09/964,824
 ; PRIOR FILING DATE: 2001-09-27
 ; PRIOR APPLICATION NUMBER: US/09/967,768
 ; PRIOR FILING DATE: 2001-09-28
 ; PRIOR APPLICATION NUMBER: US/09/968,007
 ; PRIOR FILING DATE: 2001-10-02
 ; PRIOR APPLICATION NUMBER: US/09/969,347
 ; PRIOR FILING DATE: 2001-10-02
 ; PRIOR APPLICATION NUMBER: US/09/969,708
 ; PRIOR FILING DATE: 2001-10-03
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 8447
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 4196
 ; LENGTH: 2500
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-843-641A-4196

Query Match 54.9%; Score 578.4; DB 21; Length 2500;
 Best Local Similarity 75.0%; Pred. No. 1.1e-156;
 Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;
 QY 89 ATGAAGAGCCACTCTTAAGAAAGAGTTCTCGCGGTTTGTCTATCTTTCCAAATCCAGTACC 148
 Db 397 AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTTGTCTATCTTTCCCAATCCAGTACC 456
 QY 149 CTGATATTGGAAATGTAACAGGACAGGCTTCTCTCGGACAGCAGAGAGAGGTG 208
 Db 457 ATGATATCTGGCAGATGTATAAGAGGAGAGGCTTCTTTTGGACCGCGGAGGAGGTG 516
 QY 209 ACTTATCAAGAGATCTCCCTCACTGGAAACAAAGCTTAAAGCAGATGAGAAGTACTTCTATCT 268
 Db 517 ACTCTCCAAGGACATTCAGCACTGGGAATCCCTGAACCCGAGGAGAGATATTTATAT 576
 QY 269 CTCACATCTTAGCTTTTTCAGCCAGTGTGGAATGTAATGAAATTTGGTGGAGC 328
 Db 577 CCCATGTTCTGCTTTCTTTTGCAGCAAGCGATGGCATAGTAATGAAACTTGGTGGAGC 636
 QY 329 GCTTTAGTCAGAGGTGCGAGTTCCAGAGGCTGCTGTTTCTATGGCTTCAAAATCTCA 388
 Db 637 GATTTAGCCAAAGAGTTTCAGATTACAGAAGCCCGCTGTTTCTATGGCTTCCAAATGGCCA 696
 QY 389 TCGAAGATGTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAGATCCCA 448

Db 697 TGAAGAACATACATTTCTGAATGTATAGTCTTCTTATGTGACACTTACATATAAGATCCCA 756
 QY 449 AGAAAAAGGGAATTTTATTTTAATGCAATTGAAACCATGCCCTATGTTTAAGAAAAAGCAG 508
 Db 757 AAGAAAGGGAATTTCTCTTCAATGCCATTGAAACGATGCCCTTGTGTCAAGAAAGAGCAG 816
 QY 509 ATTGGGCTTGGCATGAGATAGCAGATAAGAAATCTACTTTTGGGGAAGAGTGGTGGCCT 568
 Db 817 ACTGGGCTTGGCTGATTTGGGACAAAGAGGCTACTATGTTGTAACGCTGTTGTAGCCT 876
 QY 569 TTGCTGCTGTAGAAGGAGTTTCTTCTCAGATCTTTTGTGCTATATTTCTGGCTTAAGA 628
 Db 877 TTGCTGAGTGAAGGAGCAATTTCTTTCCGGTCTTTTGGCTCGATATTTCTGGCTCAAGA 936
 QY 629 AGAGAGTCTTATGCGCAGGACTCACTTTTCCAAATGAATCATCAGCAGAGATGAAGGAC 688
 Db 937 AAGAGGACTGATGCCCTGGCTCACATTTTCTAATGAATTTATAGCAGAGATGAGGCT 996
 QY 689 TTCACTGTGACTTTGCTTGGCTGATGTTCGAATCTTAGTAAATAAGCCTTCAGAAGAAA 748
 Db 997 TACACTGTGATTTTGGCTTGGCTGATGTTCGAATCTTAGTAAATAAGCCTTCAGAAGAAA 1056
 QY 749 GGTGAGGAGATCATTTGTTGATGCTCTCAAAATTTGAGCAGGAGCTTTTAAACAGAGCCT 808
 Db 1057 GAGTAAGAGAAATAATTTATCAATGCTGTTCCGATAGAAACAGGAGTTCTCTCAGGCT 1116
 QY 809 TGCAGTTGGCTCATTTGGAATGAATTTGATGAAACAGTACATTCAGTGTGTAG 868
 Db 1117 TGCCTGTGAAGCTCATTTGGGATGATTTGCACTCTAATGAAGCAATACATTCAGTGTGTAG 1176
 QY 869 CTGACAGATTTACTTTGTGGAATCTTGGATTTCTCAAGGTTTTCAGGCAAGAAATCTTTTG 928
 Db 1177 CAGACAGACTTATGCTGGAATCTTGGTTTTCAGAGGTTTTCAGAGTAGAGAACCCATTTG 1236
 QY 929 ATTTTATGGAACATTTCTTTAGAGGAAAAACAAATTTCTTTGAGNAACGAGTTTCAG 988
 Db 1237 ACTTTATGGAATAATTTTCACTGGAAGGAAAGACTAACTTCTTTGAGAAGAGAGTAGGCG 1296
 QY 989 AGTATCAGCGTTTTCAGGTTTATGCGAGAAACCAAGATAAGCTTCTACCTTGGATGCG 1048
 Db 1297 AGTATCAGAGGATGGGAGTGAATGTCAGTCCAAAGAGATTTCTTTTACCTTGGATGCTG 1356
 QY 1049 ATTT 1052
 Db 1357 ACTT 1360

Search completed: June 16, 2005, 11:29:37
 Job time : 756 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2005, 06:56:11 ; Search time 4079 Seconds
(without alignments)
9826.349 Million cell updates/sec

Title: US-10-698-228-2
Perfect score: 1053
Sequence: 1 atggggcaccggaaggcc.....tcaccttgatgcagatttt 1053

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1053	100.0	3292	3	CR617553 full-leng
2	1053	100.0	4748	3	AL137348 Homo sapi
3	1041	98.9	4650	3	BC042948 Homo sapi
4	1010.6	96.0	3501	3	CR627376 Homo sapi
5	945	89.7	1056	9	AY398973 Homo sapi
6	929.6	88.3	1083	1	AL547501 AL547501
7	829.8	78.8	1008	9	AY398974 Pan trogl
8	793.2	75.3	879	5	BQ441857 AGENCOURT
9	777	73.8	1043	9	AY398975 Mus muscu
10	731.2	69.4	892	7	CN163214 952695 MA
11	728.2	69.2	914	7	CF995079 AGENCOURT
12	710.6	67.5	870	7	CO648380 ILLUMIGEN
13	682.8	64.8	804	6	CD656661 AGENCOURT
14	655.4	62.2	1061	4	BM801298 AGENCOURT
15	654.8	62.2	896	5	BUI70979 AGENCOURT
16	616.4	58.5	685	5	BX951374 AGENCOURT
17	578.4	54.9	1571	3	CR603461 full-leng
18	578.4	54.9	1573	3	CR625489 full-leng
19	578.4	54.9	1582	3	CR608076 full-leng
20	578.4	54.9	1588	3	CR602054 full-leng
21	578.4	54.9	1592	3	CR621427 full-leng
22	578.4	54.9	1600	3	CR604378 full-leng
23	578.4	54.9	1605	3	CR590959 full-leng
24	578.4	54.9	1612	3	CR614990 full-leng

25	578.4	54.9	1613	3	CR609838 full-leng
26	578.4	54.9	1623	3	CR618451 full-leng
27	578.4	54.9	1630	3	CR602150 full-leng
28	578.4	54.9	1796	3	CR596700 full-leng
29	570.4	54.2	2088	3	AK088907 Mus muscu
30	562	53.4	638	7	CN358022 170005831
31	547	51.9	547	7	CN358025 170005326
32	530.8	50.4	836	7	CR416843 CR416843
33	528.8	50.2	608	1	AUI37050 AUI37050
34	527.2	50.1	540	5	BU429508 UI-HF-BNO
35	525	49.9	997	4	BM468712 AGENCOURT
36	518.8	49.3	522	5	BQ694828 1000971 H
37	513.6	48.8	793	7	CK778582 965745 MA
38	503.8	47.8	923	7	CR580780 CR580780
39	503	47.8	792	5	BU328458 603494277
40	502	47.7	510	7	CR540284 DKFZp459N
41	501.4	47.6	697	6	CA328968 UI-M-PVO-
42	493.8	46.9	510	5	EX955728 DKFZp781G
43	493.2	46.8	946	7	CN024619 AGENCOURT
44	491.8	46.7	932	5	BUI96941 AGENCOURT
45	489.4	46.5	909	6	CA981614 AGENCOURT

ALIGNMENTS

RESULT 1
CR617553 3292 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CSODI011YF14 of Placenta Cot 25-normalized
of Homo sapiens (human).
ACCESSION CR617553
VERSION CR617553.1 GI:50498360
KEYWORDS HTC; CNSLT_CDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3292)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue
REFERENCE 2 (bases 1 to 3292)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES Location/Qualifiers
source
1. .3292
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI011YF14"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 100.0%; Score 1053; DB 3; Length 3292;
Best Local Similarity 100.0%; Pred.No. 3.3e-269; Indels 0; Gaps 0;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGCGGACCGGAAAGCGGCGGCGGCTGGATCAGGATGAGATCATCTTCA 60
|||||
Db 27 ATGGCGGACCGGAAAGCGGCGGCGGCTGGATCAGGATGAGATCATCTTCA 86

Qy	61	GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC	120
Db	87	GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC	146
Qy	121	CGGTTTGTCTATCTTCCATCCAGTACCTCGATATTTGGAAATGTATATAACAGGACAG	180
Db	147	CGGTTTGTCTATCTTCCATCCAGTACCTCGATATTTGGAAATGTATATAACAGGACAG	206
Qy	181	GCCTTCTCTGGACAGCAGAGAGGTGCGATTCAAGAGATCTCCCTCACTGGAACAAG	240
Db	207	GCTTCTCTCTGGACAGCAGAGAGGTGCGATTCAAGAGATCTCCCTCACTGGAACAAG	266
Qy	241	CTTAAAGCAGATGAGAGTACTTCTCATCTCTACATCTTAGCCCTTTTTCGAGCCAGTGAT	300
Db	267	CTTAAAGCAGATGAGAGTACTTCTCATCTCTACATCTTAGCCCTTTTTCGAGCCAGTGAT	326
Qy	301	GGAAATGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGAGTGCGAGTTCCAGAGCT	360
Db	327	GGAAATGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGAGTGCGAGTTCCAGAGCT	386
Qy	361	CGCTGTTCTATGCTTCAAAATCTCATCGAATGTTCACTCAGAGATGTACAGTTTG	420
Db	387	CGCTGTTCTATGCTTCAAAATCTCATCGAATGTTCACTCAGAGATGTACAGTTTG	446
Qy	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTCGA	480
Db	447	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTCGA	506
Qy	481	ACCATGCCCTATGTTAAGAAAAAGCAGATGGGCCCTTGGATGATGATGATGATGATGATG	540
Db	507	ACCATGCCCTATGTTAAGAAAAAGCAGATGGGCCCTTGGATGATGATGATGATGATGATG	566
Qy	541	TCTACTTTTGGGAAAGAGTGGGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	600
Db	567	TCTACTTTTGGGAAAGAGTGGGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	626
Qy	601	TCTTTTCTGCTATATCTGCTTAAAGAAAGAGAGTCTTATGCCAGGACTCACCTTTTTC	660
Db	627	TCTTTTCTGCTATATCTGCTTAAAGAAAGAGAGTCTTATGCCAGGACTCACCTTTTTC	686
Qy	661	AATGAATCATCAGCAGATGAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA	720
Db	687	AATGAATCATCAGCAGATGAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA	746
Qy	721	TACTTAGTAATAAGCCTTCAAGAAAGGTCAGGAGATCATTTGATGCTGTCAAA	780
Db	747	TACTTAGTAATAAGCCTTCAAGAAAGGTCAGGAGATCATTTGATGCTGTCAAA	806
Qy	781	ATTGAGCAGAGTTTAAACAGAGCCTTGGCAGTGGCCTCATTTGGAATGAATTCGATT	840
Db	807	ATTGAGCAGAGTTTAAACAGAGCCTTGGCAGTGGCCTCATTTGGAATGAATTCGATT	866
Qy	841	TTGATGAAACAGTACATGATGTTGTAGCTGACAGATTACTTTGGAACCTTGGATTCTCA	900
Db	867	TTGATGAAACAGTACATGATGTTGTAGCTGACAGATTACTTTGGAACCTTGGATTCTCA	926
Qy	901	AAGTTTTCAGGAGAAATCCTTTTGTATTTTATGGAACATTTCTTTAGAGGAAAA	960
Db	927	AAGTTTTCAGGAGAAATCCTTTTGTATTTTATGGAACATTTCTTTAGAGGAAAA	986
Qy	961	ACAAATTTCTTGGAAACAGTTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC	1020
Db	987	ACAAATTTCTTGGAAACAGTTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC	1046
Qy	1021	ACAGATAACGCTTTCACCTTGGATGCAATTTT	1053
Db	1047	ACAGATAACGCTTTCACCTTGGATGCAATTTT	1079

RESULT 2
HSM802033
LOCUS

HSM802033 4748 bp mRNA linear HTC 22-SEP-2004

DEFINITION	Homo sapiens mRNA; cDNA DKFZp761E1312 (from clone DKFZp761E1312).
ACCESSION	AL137348
VERSION	AL137348.1 GI:6807859
KEYWORDS	HTC.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	1 (bases 1 to 4748) Ansoorge, W., Krieger, S., Regiert, T., Rittmueller, C., Schwager, B., Mewes, H.W., Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and Wiemann, S.
AUTHORS	The German cDNA Consortium
CONSTRM	Direct Submission
TITLE	Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764
JOURNAL	Neuberberg, GERMANY
COMMENT	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp761E1312) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering: http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp761E1312 Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.
FEATURES	Location/Qualifiers
source	1..4748 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="RZPD:DKFZp761E1312" /db_xref="taxon:9606" /clone="DKFZp761E1312" /tissue_type="amygdala" /clone_lib="761 (synonym: hamy2). Vector pSport1; host DH10B; sites NotI + SalI" /dev_stage="adult" /note="ribonucleotide reductase"
gene	1..4748 /gene="DKFZp761E1312"
CDS	47..1102 /gene="DKFZp761E1312" /codon_start=1 /product="hypothetical protein" /protein_id="CAB70703.2" /db_xref="GI:52545612" /db_xref="GOA:Q9NTD8" /db_xref="UniProt/TREMBL:Q9NTD8" /translation="MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLIRKSSRRFVI PFIQYDPIWMYKQASFTAEVDSKDLPHWNKLKADKDYFTSHLLAPFADSGI VNEVLVERFQSVQVPEARCFYQILLIENVHSEMYSLIDTYIRDPKKRFLFNALIE TMPYKKKADWLRADIKSTFGFVVAFAVEGVPFSGPAAIFWLKRGRLMPLGT FSNELISDEGLHCDFACLMFPQYLVNKPSEERREIIVDAVKEIEFLEALPVLG MNCILMKQYIEFVADRLVLLVGLFSGVQFAENPFDPFMENTISLEGKTNFPEKRYSEYQRF AVMAETTDNVFTLDADF"
ORIGIN	
Query Match	100.0%; Score 1053; DB 3; Length 4748;
Best Local Similarity	100.0%; Pred. No. 3.7e-269;
Matches 1053; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 ATGGCGGACCGGAAAGCGCGGAGCGCGGCTGGATCAGATGAGAGATCATCTTCA 60
Db	47 ATGGCGGACCGGAAAGCGCGGAGCGCGGCTGGATCAGATGAGAGATCATCTTCA 106
Qy	61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
Db	107 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 166
Qy	121 CGGTTTGTCTATCTTTCGAATCCAGTACCTGATATTTGGAATATGTATATAACAGGACAG 180
Db	167 CGGTTTGTCTATCTTTCGAATCCAGTACCTGATATTTGGAATATGTATATAACAGGACAG 226

181 GCTTCCTTCTGACAGCAGAGAGTGCATTTATCAAAAGATCTCCCTCAGTGGAAACAG 240
 227 GCTTCCTTCTGACAGCAGAGAGTGCATTTATCAAAAGATCTCCCTCAGTGGAAACAG 286
 241 CTTAAAGCAGATGAGAGTACTTCTCATCTCAGATCTTGTAGCTTTTTCGACGCAAGTAT 300
 287 CTTAAAGCAGATGAGAGTACTTCTCATCTCAGATCTTGTAGCTTTTTCGACGCAAGTAT 346
 301 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGAGGTGCAGGTTCCAGAGGCT 360
 347 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGAGGTGCAGGTTCCAGAGGCT 406
 361 CGCTGTTTCTATGGCTTTCAATTTCTCATCGAAGTGTCTACTCAGAGATGTACAGTTTG 420
 407 CGCTGTTTCTATGGCTTTCAATTTCTCATCGAAGTGTCTACTCAGAGATGTACAGTTTG 466
 421 CTGATAGACACTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 480
 467 CTGATAGACACTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 526
 481 ACCATGCCCTATGTTAAGAAAAAGCAGATGGGCTTGGCATGGATGATAGCAGATAGAAA 540
 527 ACCATGCCCTATGTTAAGAAAAAGCAGATGGGCTTGGCATGGATGATAGCAGATAGAAA 586
 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTCTGCTGTAGAGAGGATTTTCTTCTCAGGA 600
 587 TCTACTTTTGGGAAAGAGTGGTGGCTTTCTGCTGTAGAGAGGATTTTCTTCTCAGGA 646
 601 TCTTTTGTCTATATTTCTGCTAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
 647 TCTTTTGTCTATATTTCTGCTAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTTC 706
 661 AATGAACCTATCAGCAGAGATGAGGACTTCACTGTGATCTTGTGCTGCTGATGTTCAA 720
 707 AATGAACCTATCAGCAGAGATGAGGACTTCACTGTGATCTTGTGCTGCTGATGTTCAA 766
 721 TACTTAGTAATAAGCCTTTCAGAAAGAGGTGAGGAGATCATTTGTGATGCTGTCAA 780
 767 TACTTAGTAATAAGCCTTTCAGAAAGAGGTGAGGAGATCATTTGTGATGCTGTCAA 826
 781 ATTGAGCAGGAGTTTAAACAGAGCCCTGCGAGTGGCCTCATTTGGAATGAATGCAAT 840
 827 ATTGAGCAGGAGTTTAAACAGAGCCCTGCGAGTGGCCTCATTTGGAATGAATGCAAT 886
 841 TTGATGAAACAGTACATTTGATGCTGACAGATTTACTTGGAACTTGGATTTCA 900
 887 TTGATGAAACAGTACATTTGATGCTGACAGATTTACTTGGAACTTGGATTTCA 946
 901 AAGGTTTTCAGGCAGAGAAATCCTTTGATTTTATGAAACATTTCTTTAGAGGAAAA 960
 947 AAGGTTTTCAGGCAGAGAAATCCTTTGATTTTATGAAACATTTCTTTAGAGGAAAA 1006
 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGGTTATGCGAGAAAC 1020
 1007 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGGTTATGCGAGAAAC 1066
 1021 ACAGATACGCTTTCACCTTGGATGCAATTTT 1053
 1067 ACAGATACGCTTTCACCTTGGATGCAATTTT 1099

RESULT 3
 BC042948
 LOCUS Homo sapiens cDNA clone IMAGE:4798175, containing frame-shift errors.
 DEFINITION
 ACCESSION BC042948
 VERSION BC042948.2 GI:34194000
 KEYWORDS HTC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheet, T.B., Brownstein, M.J., Ustin, T.B., Loquellano, N.A., Peters, G.J.,
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullaby, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, R.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S.S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smallos, D.E.,
 Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL
PUBMED

12477932
 2 (bases 1 to 4650)
 Director MGC Project.
 Direct Submission
 Submitted (09-JAN-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REFERENCE
AUTHORS

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 On Aug 25, 2003 this sequence version replaced gi:27695575.
 Contact: MGC help desk
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@bcgsc.bc.ca

JOURNAL
PUBMED

12477932
 2 (bases 1 to 4650)
 Director MGC Project.
 Direct Submission
 Submitted (09-JAN-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 On Aug 25, 2003 this sequence version replaced gi:27695575.
 Contact: MGC help desk
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@bcgsc.bc.ca

FEATURES
source

1. 4650
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4798175"
 /tissue_type="Brain, hypothalamus"
 /clone_lib="NIH MGC_96"
 /lab_host="DH10B"
 /note="Vector: pBluescript"

FEATURES
Location/Qualifiers

1. 4650
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4798175"
 /tissue_type="Brain, hypothalamus"
 /clone_lib="NIH MGC_96"
 /lab_host="DH10B"
 /note="Vector: pBluescript"

FEATURES
Location/Qualifiers

1. 4650
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4798175"
 /tissue_type="Brain, hypothalamus"
 /clone_lib="NIH MGC_96"
 /lab_host="DH10B"
 /note="Vector: pBluescript"

FEATURES
Location/Qualifiers

1. 4650
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4798175"
 /tissue_type="Brain, hypothalamus"
 /clone_lib="NIH MGC_96"
 /lab_host="DH10B"
 /note="Vector: pBluescript"

FEATURES
Location/Qualifiers

1. 4650
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4798175"
 /tissue_type="Brain, hypothalamus"
 /clone_lib="NIH MGC_96"
 /lab_host="DH10B"
 /note="Vector: pBluescript"

FEATURES
Location/Qualifiers

1. 4650
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4798175"
 /tissue_type="Brain, hypothalamus"
 /clone_lib="NIH MGC_96"
 /lab_host="DH10B"
 /note="Vector: pBluescript"

FEATURES
Location/Qualifiers

1. 4650
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4798175"
 /tissue_type="Brain, hypothalamus"
 /clone_lib="NIH MGC_96"
 /lab_host="DH10B"
 /note="Vector: pBluescript"

FEATURES
Location/Qualifiers

1. 4650
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4798175"
 /tissue_type="Brain, hypothalamus"
 /clone_lib="NIH MGC_96"
 /lab_host="DH10B"
 /note="Vector: pBluescript"

FEATURES
Location/Qualifiers

1. 4650
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4798175"
 /tissue_type="Brain, hypothalamus"
 /clone_lib="NIH MGC_96"
 /lab_host="DH10B"
 /note="Vector: pBluescript"

FEATURES
Location/Qualifiers

1. 4650
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4798175"
 /tissue_type="Brain, hypothalamus"
 /clone_lib="NIH MGC_96"
 /lab_host="DH10B"
 /note="Vector: pBluescript"

FEATURES
Location/Qualifiers

1. 4650
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4798175"
 /tissue_type="Brain, hypothalamus"
 /clone_lib="NIH MGC_96"
 /lab_host="DH10B"
 /note="Vector: pBluescript"

FEATURES
Location/Qualifiers

1. 4650
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4798175"
 /tissue_type="Brain, hypothalamus"
 /clone_lib="NIH MGC_96"
 /lab_host="DH10B"
 /note="Vector: pBluescript"

FEATURES
Location/Qualifiers

1. 4650
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4798175"
 /tissue_type="Brain, hypothalamus"
 /clone_lib="NIH MGC_96"
 /lab_host="DH10B"
 /note="Vector: pBluescript"

FEATURES
Location/Qualifiers

1. 4650
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4798175"
 /tissue_type="Brain, hypothalamus"
 /clone_lib="NIH MGC_96"
 /lab_host="DH10B"
 /note="Vector: pBluescript"

FEATURES
Location/Qualifiers

1. 4650
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4798175"
 /tissue_type="Brain, hypothalamus"
 /clone_lib="NIH MGC_96"
 /lab_host="DH10B"
 /note="Vector: pBluescript"

FEATURES
Location/Qualifiers

1. 4650
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4798175"
 /tissue_type="Brain, hypothalamus"
 /clone_lib="NIH MGC_96"
 /lab_host="DH10B"
 /note="Vector: pBluescript"

FEATURES
Location/Qualifiers

1. 4650
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4798175"
 /tissue_type="Brain, hypothalamus"
 /clone_lib="NIH MGC_96"
 /lab_host="DH10B"
 /note="Vector: pBluescript"

FEATURES
Location/Qualifiers

1. 4650
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4798175"
 /tissue_type="Brain, hypothalamus"
 /clone_lib="NIH MGC_96"
 /lab_host="DH10B"
 /note="Vector: pBluescript"

FEATURES
Location/Qualifiers

1. 4650
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4798175"
 /tissue_type="Brain, hypothalamus"
 /clone_lib="NIH MGC_96"
 /lab_host="DH10B"
 /note="Vector: pBluescript"

FEATURES
Location/Qualifiers

1. 4650
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4798175"
 /tissue_type="Brain, hypothalamus"
 /clone_lib="NIH MGC_96"
 /lab_host="DH10B"
 /note="Vector: pBluescript"

FEATURES
Location/Qualifiers

1. 4650
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4798175"
 /tissue_type="Brain, hypothalamus"
 /clone_lib="NIH MGC_96"
 /lab_host="DH10B"
 /note="Vector: pBluescript"

Db 72 ATGGCGACCGGAAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 131
Qy 61 GACACCAACGAAAGTGAATAAGTCAATGAAGAGCCACTCTCAAGAAAGAGTTCTCGC 120
Db 132 GACACCAACGAAAGTGAATAAGTCAATGAAGAG-CACTCTCAAGAAAGAGTTCTCGC 190
Qy 121 CGTTTGTCTATCTTCCAAATCCAGTACCTGATATTTGGAAATGTATAACAGGCACAG 180
Db 191 CGTTTGTCTATCTTCCAAATCCAGTACCTGATATTTGGAAATGTATAACAGGCACAG 250
Qy 181 GCTTCTCTTGGACAGAGAGGTGAGCTTATCAAGAGATCTCCCTCACTGGAACAAG 240
Db 251 GCTTCTCTTGGACAGAGAGGTGAGCTTATCAAGAGATCTCCCTCACTGGAACAAG 310
Qy 241 CTTAAAGCAGATGAGAGTACTTCACTCTCACATCTTAGCTTTTGGCAGCCAGTGAT 300
Db 311 CTTAAAGCAGATGAGAGTACTTCACTCTCACATCTTAGCTTTTGGCAGCCAGTGAT 370
Qy 301 GGAATTGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTTCCAGAGGCT 360
Db 371 GGAATTGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTTCCAGAGGCT 430
Qy 361 CGCTGTTTCTATGCTTCAAAATCTCATCGAGATGTTCACTCAGAGATGTACAGTTG 420
Db 431 CGCTGTTTCTATGCTTCAAAATCTCATCGAGATGTTCACTCAGAGATGTACAGTTG 490
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 480
Db 491 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 550
Qy 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCCCTTGGATGGATGACAGATAGAAAA 540
Db 551 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCCCTTGGATGGATGACAGATAGAAAA 610
Qy 541 TCTACTTTTGGGAAAGAGTGGGCCCTTCTGCTGTAGAGAGGTTTCTTCTCAGGA 600
Db 611 TCTACTTTTGGGAAAGAGTGGGCCCTTCTGCTGTAGAGAGGTTTCTTCTCAGGA 670
Qy 601 TCTTTTCTGCTATTTCTCGCTAAAGAGAGGCTTATGCCAGGACTCACATTTTCC 660
Db 671 TCTTTTCTGCTATTTCTCGCTAAAGAGAGGCTTATGCCAGGACTCACATTTTCC 730
Qy 661 AATGAATCTATCAGCAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 731 AATGAATCTATCAGCAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 790
Qy 721 TACTTAGTAATAAGCCTTCAGAAAGAGGTCAAGGAGATCATTTGTGATGCTGTCAAA 780
Db 791 TACTTAGTAATAAGCCTTCAGAAAGAGGTCAAGGAGATCATTTGTGATGCTGTCAAA 850
Qy 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTTCGCTCATTTGGAATGCAATTGCATT 840
Db 851 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTTCGCTCATTTGGAATGCAATTGCATT 910
Qy 841 TTGATGAAACAGTACATTGATTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
Db 911 TTGATGAAACAGTACATTGATTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 970
Qy 901 AAGTTTTTTCAGGAGAAAAATCCTTTTGAATTTATGGAACAATTTCTTTAGAAAGGAAAA 960
Db 971 AAGTTTTTTCAGGAGAAAAATCCTTTTGAATTTATGGAACAATTTCTTTAGAAAGGAAAA 1030
Qy 961 ACAAATTTCTTTGAGAAACAGTTTTCAGAGTATCAGCGTTTTCAGGTTATGGCAGAACCC 1020
Db 1031 ACAAATTTCTTTGAGAAACAGTTTTCAGAGTATCAGCGTTTTCAGGTTATGGCAGAACCC 1090
Qy 1021 ACAGATAAGTCTTCACTCTGGATGCAGATTTT 1053
Db 1091 ACAGATAAGTCTTCACTCTGGATGCAGATTTT 1123

LOCUS CR627376 3501 bp mRNA linear HTC 22-SEP-2004
DEFINITION Homo sapiens mRNA; cDNA DKFZp686M05248 (from clone DKFZp686M05248).
ACCESSION CR627376
VERSION CR627376.1 GI:50949847
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 3501)
AUTHORS Ottenwälder,B.; Obermaier,B., Deutschenbaur,S., Schaipp,A.,
Mewes,H.W., Weil,B., Amlid,C., Osanger,A., Fobo,G., Han,M. and
Wiemann,S.
CONSTRM The German cDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764
COMMENT Neuberberg, GERMANY
Research from S. Wilmann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wilmann@dkfz-heidelberg.de;
sequenced by Medigenomix (Martinsried/Germany) within the cDNA
sequencing consortium of the German Genome Project.
This clone (DKFZp686M05248) is available at the RZPD Deutsches
Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp686M05248
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.
FEATURES
Location/Qualifiers
1..3501
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686M05248"
/tissue_type="salivary gland"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
/note="ribonucleotide reductase M2 B (TP53 inducible)"
gene
1..3501
/gene="DKFZp686M05248"
332..1225
/gene="DKFZp686M05248"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH10473.1"
/db_xref="GI:50949848"
/translation="MYKQASFTWAEVDLSKDLPHNKLKADKEYFISHILAPFAA
SDGTVNENLVERFSQEVQVPEARCFYQILNIENVHSEMYSLIDITYIRPKKREPLF
NAITMPYVKKADWALWIAIDRKSTFGERVVFAAAGVFFSGSFAAIFWLKKGLM
PGLTFPSNELISRDEGLHCDLAFQYLVNKPPEERVREIIVDAVKIEQEFTEALPV
GLIGNCILMKQYIEFVADRLVLELGFQVQAEVDFMENISLEGKTNPFKEKRVSE
YQRFVAVAEITDNTVFLDADF"
ORIGIN
Query Match 96.0%; Score 1010.6; DB 3; Length 3501;
Best Local Similarity 98.6%; Pred. No. 6.8e-258;
Matches 1019; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
Qy 21 GGAAGCGCCCGGCTGGATCAGGATGAGAGATCATCTTCAGACACCAACGAAGTGAAT 80
Db 190 GGAAGCGCTTGGAGGCTTCAGGCCGGGAGATCATCTTCAGACACCAACGAAGTGAAT 249
Qy 81 AAAGTCAAAATCAAGAGCCACTCTCAAGAAAGAGTTCTCGCCGTTGTCTATCTTCCAAT 140
Db 250 AAAGTCAAAATCAAGAGCCACTCTCAAGAAAGAGTTCTCGCCGTTGTCTATCTTCCAAT 309
Qy 141 CCAGTACCCTCATATTTGGAAAATGTATAAACAGGCACACAGGCTTCTTCTGGACAGCAGA 200
Db 310 CCAGTACCCTCATATTTGGAAAATGTATAAACAGGCACACAGGCTTCTTCTGGACAGCAGA 369
Qy 201 AGAGTCCAGTTATCAAAAGGATCTCCCTCACTGGAAACAAGCTTAAAGCAGATGAGAAGTA 260
Db 370 AGAGTCCAGTTATCAAAAGGATCTCCCTCACTGGAAACAAGCTTAAAGCAGATGAGAAGTA 429

QY 661 AATGAATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
 Db 661 AATGAATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
 QY 721 TACTTAGTAATAAGCCCTTCAGAGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAAA 780
 Db 721 TACTTAGTAATAAGCCCTTCAGAGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAAA 780
 QY 781 ATTGAGCAGGAGTTTAAACAGAGAGCTTCCAGTTGCGCTCATTTGGAATGAATTGCATT 840
 Db 781 ATTGAGCAGGNNNNNNNNNNNAGCCCTTCCAGTTGCGCTCATTTGGAATGAATTGCATT 840
 QY 841 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTAATTGTTGGAACCTTGGATTTCTCA 900
 Db 841 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTAATTGTTGGAACCTTGGATTTCTCA 900
 QY 901 AAGGTTTTTTCAGGAGAAATTCCTTTTGAATTTATGGAACATTTCTTTAGAGGAAAA 960
 Db 901 AAGGTTTTTTCAGGAGAAATTCCTTTTGAATTTATGGAACATTTCTTTAGAGGAAAA 960
 QY 961 ACAAATTTCTTTCAGAGAACAGTTTCAGAGTATCAGCGTTTTCAGGTTATGGCAGAAACC 1020
 Db 961 ACAAATTTCTTTCAGAGAACAGTTTCAGAGTATCAGCGTTTTCAGGTTATGGCAGAAACC 1020
 QY 1021 ACAGATAACGTTCTTACCTTTGGATGCGAGATTTT 1053
 Db 1021 ACAGATAACGTTCTTACCTTTGGATGCGAGATTTT 1053

RESULT 6

AL547501 1083 bp mRNA linear EST 25-MAR-2004
 LOCUS AL547501 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 DEFINITION clone CS0D1011YF14 5-PRIME, mRNA sequence.
 ACCESSION AL547501
 VERSION AL547501.3 GI:45747948
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1083)
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On Feb 15, 2001 this sequence version replaced gi:31269332.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 4436.r

For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?c=CS0D1011YF14&c=4436.r.

FEATURES

source
 1. .1083
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D1011YF14"
 /tissue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 88.3%; Score 929.6; DB 1; Length 1083;

Best Local Similarity 98.0%; Pred. No. 1.9e-236;
 Matches 976; Conservative 6; Mismatches 10; Indels 4; Gaps 4;
 QY 1 ATGGCGGACCCGCGAAAGCGCGCGCGG-GCTGGATCAGGATGAGAGATCATCTTTC 59
 Db 27 ATGGCGGACCCGCGAAAGCGCGCGCGGCGCGGCTGGATCAGGATGAGAGATCATCTTC 86
 QY 60 AGACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCG 119
 Db 87 AGACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCG 146
 QY 120 CCGGTTTGTCAATCTTTCCATCCAGTACCTCGATATTTTGGAAAAATGATAAAACAGGCACA 179
 Db 147 CCGGTTTGTCAATCTTTCCATCCAGTACCTCGATATTTTGGAAAAATGATAAAACAGGCACA 206
 QY 180 GGCCTTCTTCGAGCAGCAGAGAGGTCGACTTATCAAGAGATCTCCCTCACTGGAAACAA 239
 Db 207 GGCCTTCTTCGAGCAGCAGAGAGGTCGACTTATCAAGAGATCTCCCTCACTGGAAACAA 266
 QY 240 GCTTAAAGCAGATGAGAGAGTACTTTCATCTCTCACATCTTTAGCCTTTTTCAGCCAGTGA 299
 Db 267 GCTTAAAGCAGATGAGAGAGTACTTTCATCTCTCACATCTTTAGCCTTTTTCAGCCAGTGA 326
 QY 300 TGGAAATTTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGC 359
 Db 327 TGGAAATTTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGC 386
 QY 360 TCGCTGTTTCTATGCTTTTCAAAATTTCTCATCGAGAAATGTTCTACTCAGAGATGTACAGTTT 419
 Db 387 TCGCTGTTTCTATGCTTTTCAAAATTTCTCATCGAGAAATGTTCTACTCAGAGATGTACAGTTT 446
 QY 420 GCTGATGACACTTTACATCAGAGATCCCAAGAAAAAGGAAATTTTATTTAAATGCAATTGA 479
 Db 447 GCTGATGACACTTTACATCAGAGATCCCAAGAAAAAGGAAATTTTATTTAAATGCAATTGA 506
 QY 480 AACCATGCCCTATGTTAAGAAAAAGAGATTGGGCTTTCGATGATGATGAGATGAGAAA 539
 Db 507 AACCATGCCCTATGTTAAGAAAAAGAGATTGGGCTTTCGATGATGATGAGATGAGAAA 566
 QY 540 ATCTACTTTTTCGGGAAAGAGTGGGCTTTCGCTCTGTAGAGAGATTTTCTTCTCAGG 599
 Db 567 ATCTACTTTTTCGGGAAAGAGTGGGCTTTCGCTCTGTAGAGAGATTTTCTTCTCAGG 626
 QY 600 ATCTTTTCTGCTATATTCTGGCTAAAGAAAGAGAGGTCTTATGCCAGGACTCACATTTTTC 659
 Db 627 ATCTTTTCTGCTATATTCTGGCTAAAGAAAGAGAGGTCTTATGCCAGGACTCACATTTTTC 686
 QY 660 CAATGAATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCA 719
 Db 687 CAATGAATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCA 746
 QY 720 ATACTTAGTAATAAGCCCTTCAGAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAA 779
 Db 747 ATACTTAGTAATAAGCCCTTCAGAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAA 806
 QY 780 AATTGAGCAGGAGTTTAAACAGAGCCCTTGCAGTGGCCCTCATTTGGAATGAATTCGAT 839
 Db 807 AATTGAGCAGGAGTTTAAACAGAGCCCTTGCAGTGGCCCTCATTTGGAATGAATTCGAT 866
 QY 840 TTTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTAATTG-TGGAACTTTGGATTTCT 898
 Db 867 TTTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTAATTG-TGGAACTTTGGATTTCT 926
 QY 899 CAAGGTTTTTTCAGGCAGAAAAATCCCTTTTGGATTTTATGAAAAACATTTCTTTAGAGGAA 958
 Db 927 CAAGGTTTTTTCAGGCAGAAAAATCCCTTTTGGATTTTATGG-AAAMATTTCTTTAGAGG-A 984
 QY 959 AAACAAATTTCTTTTCAGAGAACAGGTTTTCAGAGTATC 994
 Db 985 AAAMAAATTTCTTTTCAGAGAACAGGTTTTCAGAGTATC 1020

RESULT 7

```

AY398974      AV398974      1008 bp      DNA      linear      GSS 12-DEC-2003
LOCUS         Pan troglodytes HCM0069 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION    genomic survey sequence.
ACCESSION     AV398974
VERSION       AV398974.1 GI:39754963
KEYWORDS      GSS.
SOURCE        Pan troglodytes (chimpanzee)
ORGANISM      Pan troglodytes
REFERENCE     1 (bases 1 to 1008)
AUTHORS       Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
              Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
              Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
              Adams,M.D. and Cargill,M.
TITLE         Inferring nonneutral evolution from human-chimp-mouse orthologous
              gene trios
JOURNAL       Science 302 (5652), 1960-1963 (2003)
PUBMED       14671302
REFERENCE     2 (bases 1 to 1008)
AUTHORS       Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
              Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
              Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
              Adams,M.D. and Cargill,M.
TITLE         Direct Submission
JOURNAL       Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
              Rockville, MD 20850, USA
COMMENT       This sequence was made by sequencing genomic exons and ordering
              them based on alignment.
FEATURES      Location/Qualifiers
               source
               1..1008
               /organism="Pan troglodytes"
               /mol_type="genomic DNA"
               /db_xref="taxon:9598"
               <1..>1008
               /locus_tag="HCM0069"
ORIGIN
Query Match      78.8%; Score 829.8; DB 9; Length 1008;
Best Local Similarity 82.7%; Pred. No. 7.6e-210;
Matches 831; Conservative 0; Mismatches 174; Indels 0; Gaps 0;

Qy 49 AGATCATCTTCAGACACCAACGAAAGTGAATAAGTCAATGAAGACCACTCCTAAGA 108
Db 1 AGATCATCTTCAGACACCAACGAAAGTGAATAAGTCAATGAAGACCACTCCTAAGA 60

Qy 109 AAGAGTTCTCGCCGGTTGTTCATCTTTCCCAATCCAGTACCCTGATATTTGGAAATGTAT 168
Db 61 AAGAGTTCTCGCCGGTTGTTCATCTTTCCCAATCCAGTACCCTGATATTTGGAAATGTAT 120

Qy 169 AAACAGGACAGGCTTCCTTCGACAGCAGAGAGGTGCACTTATCAAGAGATCTCCCT 228
Db 121 AAACAGGACAGGCTTCCTTCGACAGCAGAGAGGTGCACTTATCAAGAGATCTCCCT 180

Qy 229 CACTGGACAGCTTAAAGCAGATGAGAGTACTTCTCATCTTCACATAGCCCTTTT 288
Db 181 CACTNGAACAGCTTAAAGCAGATGAGAGTACTTCTCATCTTCACATAGCCCTTTT 240

Qy 289 GCAGCCAGTGATGAAATTTGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAG 348
Db 241 GCAGCCAGTGATGAAATTTGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAG 300

Qy 349 GTTCAGAGGCTCGCTGTTTCATGGCTTCAAAATTCATCGAAGATGTTCACTCAGAG 408
Db 301 GTTCAGAGGCTCGCTGTTTCATGGCTTCAAAATTCATCGAAGATGTTCACTCAGAG 360

Qy 409 ATGTACAGTTTGTCTGATAGACACTTACATCAGATCCCAAGAAAGGGAATTTTATTT 468
Db 361 ATGTACAGTTTGTCTGATAGACACTTACATCAGATCCCAAGAAAGGGAATTTTATTT 420

Qy 469 AATGCAATTGAACACCATGCGCCCTATGTTTAAAGAAAAAGCAGATTTGGGCTTGGATGATA 528

```

```

Db 421 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 480
Qy 529 GCAGATAGAAAATCTACTTTTTGGGAAAGAGTGGTGGCCTTTGCTGCTGTAGAGAGATT 588
Db 481 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 540
Qy 589 TTCTTCTCAGGATCTTTTGTCTGCTATATTCTGGCTTAAAGAGAGAGGTTCTTATGCCAGA 648
Db 541 TTCTTCTCAGGATCTTTTGTCTGCTATATTCTGGCTTAAAGAGAGAGGTTCTTATGCCNNNN 600
Qy 649 CTCACCTTTTCCAAATGAACATCATCAGCAGATGAAGGACTTCACTGTGACATTGCTTGC 708
Db 601 NTCACCTTTTCCAAATGAACATCANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 660
Qy 709 CTGATGTTTCCAAATCTAGTAAATAAGCCTTCAGAGAAGAGGTCAGGAGAGATCATTTGTT 768
Db 661 CTGATGTTTCCAAATCTAGTAAATAAGCCTTCAGAGAAGAGGTCAGGAGAGATCATTTGTT 720
Qy 769 GATGCTGTCAAAATTTAGCAGGAGTTTAAACAGAAAGCCTTGCCAGTTGGCCTCATTTGA 828
Db 721 GATGCTGTCAAAATTTAGCAGGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTCATTTGA 780
Qy 829 ATGAATTTGCATTTTATGATAAAGTACATTTGAGTTTGTAGCTGACAGATTAATTTGTTGA 888
Db 781 ATGAATTTGCATTTTATGATGANAAGTACATTTGAGTTTGTAGCTGACAGATTAATTTGTTGA 840
Qy 889 CTGCGATTCTCAAAAGTTTTCAGGACAGAAATCCTTTGATTTTATGAAAAACATTTCT 948
Db 841 CTGCGATTCTCAAAAGTTTTCAGGACAGAAATNNNTTTGATTTTATGAAAAACATTTCT 900
Qy 949 TTAGAAGGAAAAACAAATTTCTTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGATT 1008
Db 901 TTAGAAGGAAAAACAAANNNNNNTTCAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGATT 960
Qy 1009 ATGCAGAGAACCCAGATTAAGCTTTCACCTTGGATGACAGATTTT 1053
Db 961 ATGCAGAGAACCCAGATTAAGCANNNTTCCCTTGGATGACAGATTTT 1005

RESULT 8
LOCUS    BQ441857
DEFINITION AGENCOURT 7827448 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:6014179
          5', mRNA sequence.
ACCESSION BQ441857
VERSION    BQ441857.1 GI:21180933
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 879)
AUTHORS    NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
           Email: csapbs-remail.nih.gov
           Tissue Procurement: ATCC
           CDNA Library Preparation: Life Technologies, Inc.
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Agencourt Bioscience Corporation
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LHAM3208 row: d column: 20
           High quality sequence stop: 662.

FEATURES   Location/Qualifiers
            1..879
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:6014179"
            /tissue_type="embryonal carcinoma, cell line"

```

/lab_host="DH10B (phage-resistant)" /clone_lib="NIH_MGC_92" /note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."	
ORIGIN	
Query Match	75.3%; Score 793.2; DB 5; Length 879;
Best Local Similarity	98.7%; Pred. No. 4.2e-200; Indels 3; Gaps 2;
Matches	821; Conservative 0; Mismatches 8;
Qy	1 ATGGCGCAGCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db	22 ATGGCGCAGCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 81
Qy	61 GACACCAAGCAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGTTCCTCG 120
Db	82 GACACCAAGCAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGTTCCTCG 141
Qy	121 CGGTTCGTCTCTTCAATCCAGTACCTGATATTTGGAAATGTATAACAGGCACAG 180
Db	142 CGGTTCGTCTCTTCAATCCAGTACCTGATATTTGGAAATGTATAACAGGCACAG 201
Qy	181 GCTTCCTTCTGGACAGAGAGGTGACTTATCAAAAGGATCTCCCTCACTGGAACAAG 240
Db	202 GCTTCCTTCTGGACAGAGAGGTGACTTATCAAAAGGATCTCCCTCACTGGAACAAG 261
Qy	241 CTTAAAGCAGATGAGAGTACTTCACTCTCAGATCTTACCTAGGAGGTTCAGAGGCT 300
Db	262 CTTAAAGCAGATGAGAGTACTTCACTCTCAGATCTTACCTAGGAGGTTCAGAGGCT 321
Qy	301 GGAATTCGTAATGAATTTGGTGGAGCGCTTTAGTCAGAGGTTCAGAGGCT 360
Db	322 GGAATTCGTAATGAATTTGGTGGAGCGCTTTAGTCAGAGGTTCAGAGGCT 381
Qy	361 CGCTGTTTCTATGCTTTCAAAATCTCATCAGAGATGTTCACTCAGAGATGTACAGTTTG 420
Db	382 CGCTGTTTCTATGCTTTCAAAATCTCATCAGAGATGTTCACTCAGAGATGTACAGTTTG 441
Qy	421 CTGATAGACATTAATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 480
Db	442 CTGATAGACATTAATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 501
Qy	481 ACCATGCCCTATGTTAGAAAAGACAGATTTGGCCCTTGGATGATGAGATAGAAA 540
Db	502 ACCATGCCCTATGTTAGAAAAGACAGATTTGGCCCTTGGATGATGAGATAGAAA 561
Qy	541 TCTACTTTTGGGAAAGAGTGTGGCCCTTGTCTGTAGAGAGGATTTTCTTCTCAGGA 600
Db	562 TCTACTTTTGGGAAAGAGTGTGGCCCTTGTCTGTAGAGAGGATTTTCTTCTCAGGA 621
Qy	601 TCTTTTGTCTATATTTCTGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
Db	622 TCTTTTGTCTATATTTCTGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 681
Qy	661 AATGAATCATCAGCAGATGAGAGCTTCACTGATCTTGTCTGCTGATGTTCCAA 720
Db	682 AATGAATCATCAGCAGATGAGAGCTTCACTGATCTTGTCTGCTGATGTTCCAA 741
Qy	721 TACTTTAGTAATAGCCCTTCAGAGAGAGGGG - TCAGGGAGATCATTTGTTGATGCTGTCA 778
Db	742 TACTTTAGTAATAGCCCTTCAGAGAGAGGGGTCAGGGAGATCATTTGTTGATGCTGTCA 801
Qy	779 AAATTT-GAGCAGGAGTTTTTAACAGAGAGCCCTTGCAGTGTGCCCTCATTTGAA 829
Db	802 AAATTTGAGCAGGAGTTTTTAACAGAGAGCCCTTGCAGTGTGCCCTCAATTTGA 853
RESULT 9	
AY398975	1043 bp DNA linear GSS 12-DEC-2003
LOCUS	

DEFINITION	Mus musculus HCM0069 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION	AY398975
VERSION	AY398975.1 GI:39754964
KEYWORDS	GSS.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	1 (bases 1 to 1043) Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M. Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
AUTHORS	Science 302 (5652), 1960-1963 (2003) 14671302
TITLE	2 (bases 1 to 1043) Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M. Direct Submission
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES	Location/Qualifiers source 1..1043 /organism="Mus musculus" /mol_type="genomic DNA" /db_xref="taxon:10090" <1..>1043 /locus_tag="HCM0069"
gene	
ORIGIN	
Query Match	73.8%; Score 777; DB 9; Length 1043;
Best Local Similarity	81.2%; Pred. No. 9.2e-196;
Matches	854; Conservative 0; Mismatches 185; Indels 13; Gaps 1;
Qy	1 ATGGCGCAGCCGGAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db	1 ATGGCGCAGCCGGAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Qy	61 GACACCAAGCAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGTTCCTCG 120
Db	61 GACACCAAGCAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGTTCCTCG 120
Qy	121 CGGTTCGTCTCTTCAATCCAGTACCTGATATTTGGAAATGTATAACAGGCACAG 180
Db	121 CGGTTCGTCTCTTCAATCCAGTACCTGATATTTGGAAATGTATAACAGGCACAG 180
Qy	181 GCTTCCTTCTGGACAGAGAGGTGACTTATCAAAAGGATCTCCCTCACTGGAACAAG 240
Db	181 GCTTCCTTCTGGACAGAGAGGTGACTTATCAAAAGGATCTCCCTCACTGGAACAAG 240
Qy	241 CTTAAAGCAGATGAGAGTACTTCACTCTCAGATCTTGTAGCCCTTTTTCAGCCAGTGT 300
Db	241 CTTAAAGCAGATGAGAGTACTTCACTCTCAGATCTTGTAGCCCTTTTTCAGCCAGTGT 300
Qy	301 GGAATTCGTAATGAATTTGGTGGAGCGCTTTAGTCAGAGGTTCAGAGGCT 360
Db	301 GGAATTCGTAATGAATTTGGTGGAGCGCTTTAGTCAGAGGTTCAGAGGCT 360
Qy	361 CGCTGTTTCTATGCTTTCAAAATCTCATCAGAGATGTTCACTCAGAGATGTACAGTTTG 420
Db	361 CGCTGTTTCTATGCTTTCAAAATCTCATCAGAGATGTTCACTCAGAGATGTACAGTTTG 420
Qy	421 CTGATAGACATTAATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 480
Db	421 CTGATAGACATTAATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 480


```
Db 744 CATGAGCTATCCACCAAGATGAGGAGCTTCCCTGTGAACTTTGCTTGCCTGTGCTGT 803
Qy 716 TCCAATCTTAGTAA 731
Db 804 TCCAATCTAGTAA 819

RESULT 15
BUI70979
LOCUS AGENCOURT_7940338 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6143754
DEFINITION 5', mRNA sequence.
ACCESSION BUI70979
VERSION BUI70979.1 GI:22684963
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 896)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13467 Row: k Column: 19
High quality sequence stop: 619.
FEATURES
    source
        1..896
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:6143754"
            /tissue_type="retinoblastoma"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH MGC 67"
            /notes="Organ: eye; Vector: pCMV-SPORT6; Site 1: Not I;
            Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
            Average insert size 1.75 kb. Library constructed by Life
            Technologies."
ORIGIN
    Query Match 62.2%; Score 654.8; DB 5; Length 896;
    Best Local Similarity 98.7%; Pred. No. 3.4e-163;
    Matches 671; Conservative 0; Mismatches 7; Indels 2; Gaps 1;
    Qy 1 ATGGGGACCCGGAAAGCGCGGAGCGCGGCTGGATCAGATGAGATCATCTTCA 60
    Db 55 ATGGGGACCCGGAAAGCGCGGAGCGCGGCTGGATCAGATGAGATCATCTTCA 114
    Qy 61 GACACCAACGAAAGTGAATAAATGAAGTGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
    Db 115 GACACCAACGAAAGTGAATAAATGAAGTGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 174
    Qy 121 CGCTTTGTGATCTTCCATCCAGTACCTGATATTTGGAATGATTAACAGGCACAG 180
    Db 175 CGCTTTGTGATCTTCCATCCAGTACCTGATATTTGGAATGATTAACAGGCACAG 234
    Qy 181 GCTTCTCTTGGACAGCAGAGAGGTGCACTTATCAAGAGATCTCCCTCACTGGAAACAG 240
    Db 235 GCTTCTCTTGGACAGCAGAGAGGTGCACTTATCAAGAGATCTCCCTCACTGGAAACAG 294
    Qy 241 CTTAAAGCAGATGAGAAGTACTTTTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTGAT 300
    Db 295 CTTAAAGCAGATGAGAAGTACTTTTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTGAT 354
```

```
Qy 301 GGAATGTGAATGAAAAATTTGGTGAGCGCTTTAGTCAAGAGGTGCGAGGTTCCAGAGGCT 360
Db 355 GGAATGTGAATGAAAAATTTGGTGAGCGCTTTAGTCAAGAGGTGCGAGGTTCCAGAGGCT 414
Qy 361 CGCTGTTTCTATGGCTTTCAAATTCCTCATCCAGAAATGTTCACTCAGAGATGTACAGTTTG 420
Db 415 CGCTGTTTCTATGGCTTTCAAATTCCTCATCCAGAAATGTTCACTCAGAGATGTACAGTTTG 474
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 475 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAATGCAATTGAA 534
Qy 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTTCGATGGATGAGATAGAGATAAGAAA 540
Db 535 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTTCGATGGATGAGATAGAGATAAGAAA 594
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCCCTTGTCTGTAGAGAGGTTTCTTCTCAGGA 600
Db 595 TCTACTTTTGGGAAAGAGTGGTGGCCCTTGTCTGTAGAGAGGTTTCTTCTCAGGA 654
Qy 601 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGTCTTTATGCCAGGACTCACTTTTTC 658
Db 655 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGTCTTTATGCCAGGACTCACTTTTTC 714
Qy 659 CCAATGAACTCATCAGCAGA 678
Db 715 CAATGAACTCATCAGCAGA 734
```

Search completed: June 16, 2005, 09:39:41
Job time : 4085 secs


```
Qy 261 CTTCAATCTCTCACATCTTAGCTTTTGGAGCAGTGTAGGAATGTAAATGAAAATTT 320
Db |||||
Qy 430 CTTCAATCTCTCACATCTTAGCTTTTGGAGCAGTGTAGGAATGTAAATGAAAATTT 489
Db |||||
Qy 321 GGTGGAGCGCTTTAGTCAGAGGTCAGGTTCCAGAGGCTCGCTGTCTTCTATGCTTTCA 380
Db |||||
Qy 490 GGTGGAAACGCTTTAGTCAGAGGTCAGGTTCCAGAGGCTCGCTGTCTTCTATGCTTTCA 549
Db |||||
Qy 381 AATTCTCATCAGAAATGTTCACTCAGAGATGTACAGTTTGTGTATAGACACTTACATCAG 440
Db |||||
Qy 550 AATTCTCATCAGAAATGTTCACTCAGAGATGTACAGTTTGTGTATAGACACTTACATCAG 609
Db |||||
Qy 441 AGATCCCAAGAAAAGGAAATTTTAAATGCAATTGCAATCCCAATGCTTATGTAAGAA 500
Db |||||
Qy 610 AGATCCCAAGAAAAGGAAATTTTAAATGCAATTGCAATCCCAATGCTTATGTAAGAA 669
Db |||||
Qy 501 AAAAGCAGATGGGCTTGGATGGATAGCAGATAGAAAATCTACTTTTGGGGAAGAGT 560
Db |||||
Qy 670 AAAAGCAGATGGGCTTGGATGGATAGCAGATAGAAAATCTACTTTTGGGGAAGAGT 729
Db |||||
Qy 561 GGTGGCCTTTGCTGCTGTAGAGGAGTTTCTCTCAGAGATCTTTTGTGCTGTATATCTG 620
Db |||||
Qy 730 GGTGGCCTTTGCTGCTGTAGAGGAGTTTCTCTCAGAGATCTTTTGTGCTGTATATCTG 789
Db |||||
Qy 621 GCTAAAGAAGAGAGTCTTATGCCAGGACTCACTTTTCCAAATGAACTCATCAGCAGAGA 680
Db |||||
Qy 790 GCTAAAGAAGAGAGTCTTATGCCAGGACTCACTTTTCCAAATGAACTCATCAGCAGAGA 849
Db |||||
Qy 681 TGAAGGACTTCACGTGACCTTTGCTGCTGATGTTTCAATACCTAGTAAATAGCCCTTC 740
Db |||||
Qy 850 TGAAGGACTTCACGTGACCTTTGCTGCTGATGTTTCAATACCTAGTAAATAGCCCTTC 909
Db |||||
Qy 741 AGAAGAAAGGTCAGGAGATCATTTGTGCTGTCTCAAAATTCAGCAGGAGTTTAAAC 800
Db |||||
Qy 910 AGAAGAAAGGTCAGGAGATCATTTGTGCTGTCTCAAAATTCAGCAGGAGTTTAAAC 969
Db |||||
Qy 801 AGAAGCCTTGGCAGTGGCTTCAATGGAATGAAATGCAATTTGATGAAACAGTACATTGA 860
Db |||||
Qy 970 AGAAGCCTTGGCAGTGGCTTCAATGGAATGAAATGCAATTTGATGAAACAGTACATTGA 1029
Db |||||
Qy 861 GTTTGTAGCTGACAGATTACTTTGGAACTTGGAACTTCAAGGTTTTTTCAGGCAGAAA 920
Db |||||
Qy 1030 GTTTGTAGCTGACAGATTACTTTGGAACTTGGAACTTCAAGGTTTTTTCAGGCAGAAA 1089
Db |||||
Qy 921 TCCTTTGTATTTATGAAAACATTTCTTTAGAGGAAAACAAATTTCTTTGAGAAACG 980
Db |||||
Qy 1090 TCCTTTGTATTTATGAAAACATTTCTTTAGAGGAAAACAAATTTCTTTGAGAAACG 1149
Db |||||
Qy 981 AGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACACACAGATAACGTTCTTCACTT 1040
Db |||||
Qy 1150 AGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACACACAGATAACGTTCTTCACTT 1209
Db |||||
Qy 1041 GGATGCAGATTTT 1053
Db |||||
Qy 1210 GGATGCAGATTTT 1222
Db |||||
```

RESULT 5

AY398973

LOCUS

DEFINITION Homo sapiens HCM0069 gene, VIRTUAL TRANSCRIPT, partial sequence, 1056 bp DNA linear GSS 12-DEC-2003

genomic survey sequence.

ACCESSION AY398973

VERSION AY398973.1

KEYWORDS GSS.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1056)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,

Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,

Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,

Adams,M.D. and Cargill,M.

Inferring nonneutral evolution from human-chimp-mouse orthologous

gene trios

Science 302 (5652), 1960-1963 (2003)

14671302

2 (bases 1 to 1056)

Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,

Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,

Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,

Adams,M.D. and Cargill,M.

Direct Submission

Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,

Rockville, MD 20850, USA

This sequence was made by sequencing genomic exons and ordering

them based on alignment.

FEATURES

source

1..1056

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

<1..>1056

/locus_tag="HCM0069"

ORIGIN

Query Match 89.7%; Score 945; DB 9; Length 1056;

Best Local Similarity 89.7%; Pred. No. 1.5e-240;

Matches 945; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

Qy 1 ATGGCGCACCGGAAAGCGCGGCGGCTGGATCAGATGAGATCATCTTCA 60

Db 1 ATGGCGCACCGGAAAGCGCGGCGGCTGGATCAGATGAGATCATCTTCA 60

Qy 61 GACACCAAGAAAGTGAATTAAGTCAATGAAGGCCACTCTTAGAAGAGTTCTCGC 120

Db 61 GACACCAAGAAAGTGAATTAAGTCAATGAAGGCCACTCTTAGAAGAGTTCTCGC 120

Qy 121 CGGTTTGTCTCTTTCCCAATCCAGTACCTCGATATTTGGAAAATGTATAAAGGACAG 180

Db 121 CGGTTTGTCTCTTTCCCAATCCAGTACCTCGATATTTGGAAAATGTATAAAGGACAG 180

Qy 181 GCTTCTCTTCTGGACAGCAGAGAGTTCGACTTATCAAAAGGATCTCCCTCACTGGAA 240

Db 181 GCTTCTCTTCTGGACAGCAGAGAGTTCGACTTATCAAAAGGATCTCCCTCACTGGAA 240

Qy 241 CTTAAAGCAGATGAGAAGTACTTTCATCTCTACATCTTAGCCCTTTTTCAGGCGAGT 300

Db 241 CTTAAAGCAGATGAGAAGTACTTTCATCTCTACATCTTAGCCCTTTTTCAGGCGAGT 300

Qy 301 GGAATTGTAAATGAAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360

Db 301 GGAATTGTAAATGAAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360

Qy 361 CGCTGTTTCTATGCTTTTCAAAATTTCTCATCGAATGTTTCACTCAGAGATGTACAGTTTG 420

Db 361 CGCTGTTTCTATGCTTTTCAAAATTTCTCATCGAATGTTTCACTCAGAGATGTACAGTTTG 420

Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTGAA 480

Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTGAA 480

Qy 481 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGGCTTCGATGGATAGCAGATAGAAA 540

Db 481 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 540

Qy 541 TCTACTTTTGGGGAAGAGTGTGGCTTTTGTCTGTAGAAGGAGTCTTCTTCTCAGGA 600

Db 541 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 600

Qy 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660

Db 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660

```
QY 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTGGGCTTGGATGATGATAGCAGATAGAAAA 540
Db 481 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 540
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAGGAGTGTCTCTCTCAGGA 600
Db 541 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 600
QY 601 TCTTTTGGCTATATCTGCTTAAAGAGAGAGTCTTATGCCAGGACTCAGCTTTTTC 660
Db 601 TCAATTTGCTCAATATCTGCTTAAAGAGAGAGGCTCATGCTGGACTGACTTTTTC 660
QY 661 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTGCTGATGTCCAA 720
Db 661 AATGAACATCATCAGCAGGAGTGAAGGACTTCACTGTGACTTTGCTGCTGATGTTCAG 720
QY 721 TACTTTAGTAATAAGCCTTCAGAGAGAGGCTCAGGAGAGATCATTTGTGATGCTGTC 780
Db 721 TACTTTAGTAATAAGCCTTCAGAGAGATGAGTGAAGGAAATCATTTGCTGATGCTGTC 780
QY 781 ATTGAGCAGGAGTGTAAACAGAGAGCCTTGCCAGTTGGCCTCATTTGGAATGAATTC 840
Db 781 ATCGAG-----CAGAGCCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 827
QY 841 TTGATGAACAGTACATGAGTTTGTAGTCAAGATTAATCTTTGAGAACTTTGATTTCTCA 900
Db 828 TTGATGAACAGTACATGAGTTTGTAGTCAAGATTAATCTTTGAGAACTTTGATTTCTCA 887
QY 901 AAGTTTTTTCAGGAGAGAAATCCTTTGATTTTATGGAACATTTCTTTAGAGGAAAA 960
Db 888 AAGATTTTTCAGGAGAGAAATCCTTTGATTTTATGGAACATTTCTTTAGAGGAAAA 947
QY 961 ACAAATTTCTTTGAGAAACAGTTTTCAGAGATCAGGCTTTTTCAGTTATGCGAGAAACC 1020
Db 948 ACAAATTTCTTTGAGAAACAGTTTTCAGAGATCAGGCTTTTTCAGTTATGCGAGAAACC 1007
QY 1021 ACAGATAACGCTTTCACCTTGGATGCGAGATT 1052
Db 1008 ACAGATAATGCTTTCACCTTGGATGCGAGATT 1039
```

```
RESULT 10
CN163214/c
LOCUS 952695 MARC 4PIG Sus scrofa cDNA 3', mRNA sequence. EST 02-APR-2004
ACCESSION CN163214
VERSION CN163214.1 GI:46177644
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 892)
Smith,T.P.L., Freking,B.A., Ford,J.J., Vallet,J.L., Wise,T.A.,
Nonnenman,D.J., Wray,J.E. and Keele,J.W.
Porcine EST collection using a normalized library constructed from
embryos representing early developmental stages
Unpublished (2003)
JOURNAL Contact: Smith TPL
COMMENT USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross match v0.990329.
Plate: TMW8060 row: G column: 4
Seq primer: TAGAAGCAGCAGTCGAG.
Location/Qualifiers
1. .892
/organism="Sus scrofa"
/mol_type="mRNA"
```

FEATURES
source

RESULT 11
CF995079
LOCUS

CF995079 914 bp mRNA linear EST 25-NOV-2003

/db_xref="taxon:9823"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 4PIG"
/note="Vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI;
library made with combined RNA from day-10, day-13,
day-15, day-25, and day-30 whole embryos."

ORIGIN

```
Query Match 69.4%; Score 731.2; DB 7; Length 892;  
Best Local Similarity 94.1%; Pred. No. 1.4e-183;  
Matches 760; Conservative 0; Mismatches 48; Indels 0; Gaps 0;  
QY 246 AGCAGATGAGAAGTACTTCATCTCAGATCTTGTAGTCCAGTCTTTTGGACGAGTGAAT 305  
Db 892 ATCAGATGAGAAGTATTTTATCTCTCAGATCTTGTAGTCCAGTCTTTTGGACGAGTGAAT 893  
QY 306 TGTAAATGAAAATTTGGTGGAGCGCTTTAGTCAAGAGGTGCAGGTTCCAGAGCGCTGCTG 365  
Db 832 TGTGAATGAAAATTTGGTGGAGCGCTTTAGTCAAGAGGTGCAGGTTCCAGAGCGCTGCTG 773  
QY 366 TTTCTATGCTTTCAAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTGTCTGAT 425  
Db 772 TTTCTATGCTTTCAAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTGTCTGAT 713  
QY 426 AGACACTTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTTGAAACCAT 485  
Db 712 AGACACTTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTTGAGACAAT 653  
QY 486 GCCCTATGTTAAGAAAAGAGATTTGGGCTTTGGATGGAATAGCAGATAGAAAATCTAC 545  
Db 652 GCCATATGTTAAGAAAAGAGATTTGGGCTTTGGATGGAATAGCAGATAGAAAATCTAC 593  
QY 546 TTTTGGGAAAAGAGTGGTGGCTTTTGGCTGTAGAGAGATTTTCTTCTCAGGATCTTT 605  
Db 592 TTTTGGGAAAAGAGTGGTGGCTTTTGGCTGTAGAGAGATTTTCTTCTCAGGATCTTT 533  
QY 606 TGCTGCTATATTTCTGGCTTAAAGAGAGAGGCTTTTATGCGAGACTCACATTTTCCAAATGA 665  
Db 532 TGCTGCTATATTTCTGGCTTAAAGAGAGAGGCTTTTATGCGAGACTCACATTTTCCAAATGA 473  
QY 666 ACTCATCAGCAGAGATGAAGGACTTTCATGTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCT 725  
Db 472 ACTCATCAGCAGAGATGAAGGACTTTCATGTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCT 413  
QY 726 AGTAAATAGGCTTTTCAAGAAAAGAGGCTCAGGAGATCATTTGTGCTGCTGCTGCTGCTGCT 785  
Db 412 GGTAAATAGGCTTTTCAAGAAAAGAGGCTCAGGAGATCATTTGTGCTGCTGCTGCTGCTGCT 353  
QY 786 GCAGAGTTTTAAACAGAGAGCCTTCCAGTTCGCCCTCATTTGGAATGAATTTGCAATTTGAT 845  
Db 352 GCAGAGTTTTAAACAGAGAGCCTTCCAGTTCGCCCTCATTTGGAATGAATTTGCTTTGAT 293  
QY 846 GAAACAGTATCATTTGAGTTTGTAGCTGACAGATTTACTTTGGAACCTTTGGAATTTGGAAGGT 905  
Db 292 GAAACAGTATTTGAGTTTGTAGCTGACAGATTTGCTTACAGAACTTTGGAATTTCTCAAAGGT 233  
QY 906 TTTTCAGGAGAAAATCCTTTTGTATTTTATGAAAACATTTCTTTAGAGGAAAACAAA 965  
Db 232 TTTTCAGGAGAAAATCCTTTTGTATTTTATGAAAACATTTCTTTAGAGGAAAACAAA 173  
QY 966 TTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTATGCGATGCGCAAGAACACAGA 1025  
Db 172 TTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTATGCGATGCGCAAGAACACAGA 113  
QY 1026 TAAAGCTCTTCACTTTGATGAGATTTT 1053  
Db 112 TAAAGCTCTTCACTTTGATGAGATTTT 85
```

Swape, S.
10/698228
Seq. ID 12 w/Integ

> O <
O| |o IntelliGenetics
> O <

GENALIGN - Multiple Sequence Alignment Program
Release 5.4

Mon 26 Sep 105 11:01:11-PST

Solution Parameters:

Nucleic Alphabet = Identity
Output line length = 80
Compress = Off
Histogram = Off
Randomization = Off
AMINO-Res-length = 2
Deletion-weight = 5.00
Length-factor = 0
Matching-weight = 1.00
NUCLEIC-Res-length = 4
Spread-factor = 50

Clustered order of selected sequences:

1. US-10-698-228-2 (1-1053)
10. US-10-698-228-12 (1-1053)

-Region Alignment: (listed in clustered order)

US-10-698-	1	ATGGCGACCCCGGAAGCGCGCGGCTGGATCAGGATGAGAGATCATCTTCAG
US-10-698-	1	ATGGCGACCCCGGAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCAG
consensus		ATGGCGACCCCGGAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCAG
US-10-698-	62	ACACCAAGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGCG
US-10-698-	62	ACACCAAGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGCG
consensus		ACACCAAGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGCG
US-10-698-	123	GTTTGTCTCTTCCAAATCCAGTACCTCGATATTTGGAAATGTATAACAGGCACAGGT
US-10-698-	123	GTTTGTCTCTTCCAAATCCAGTACCTCGATATTTGGAAATGTATAACAGGCACAGGT
consensus		GTTTGTCTCTTCCAAATCCAGTACCTCGATATTTGGAAATGTATAACAGGCACAGGT
US-10-698-	184	TCCTTCTGGAAGAGAGAGGTGACTTATCAAAAGGATCTCCCTCACTGGAAACAGCTTA
US-10-698-	184	TCCTTCTGGAAGAGAGAGGTGACTTATCAAAAGGATCTCCCTCACTGGAAACAGCTTA
consensus		TCCTTCTGGAAGAGAGAGGTGACTTATCAAAAGGATCTCCCTCACTGGAAACAGCTTA
US-10-698-	245	AAGCAGATGAGAAGTACTCTCATCTCATCTTAGCCCTTTTTCAGCCAGTGGATGAAT
US-10-698-	245	AAGCAGATGAGAAGTACTCTCATCTCATCTTAGCCCTTTTTCAGCCAGTGGATGAAT
consensus		AAGCAGATGAGAAGTACTCTCATCTCATCTTAGCCCTTTTTCAGCCAGTGGATGAAT
US-10-698-	306	TGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAAGGTTCCAGAGGCTCGCTGT
US-10-698-	306	TGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAAGGTTCCAGAGGCTCGCTGT
consensus		TGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAAGGTTCCAGAGGCTCGCTGT

US-10-698-	367	TTCTATGGCTTTCAAAATCTCATCGAGAAATGTTCACTCAGAGATGACAGTTTCTGCTGATAG
US-10-698-	367	TTCTATGGCTTTCAAAATCTCATCGAGAAATGTTCACTCAGAGATGACAGTTTCTGCTGATAG
consensus		TTCTATGGCTTTCAAAATCTCATCGAGAAATGTTCACTCAGAGATGACAGTTTCTGCTGATAG
US-10-698-	428	ACACTTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTTGAAACCATGCC
US-10-698-	428	ACACTTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTTGAAACCATGCC
consensus		ACACTTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTTGAAACCATGCC
US-10-698-	489	CTATGTTAAGAAAAAGCAGATTTGGGCTTTGGATGAGATAGCAGATAGAAATCTACTTTT
US-10-698-	489	CTATGTTAAGAAAAAGCAGATTTGGGCTTTGGATGAGATAGCAGATAGAAATCTACTTTT
consensus		CTATGTTAAGAAAAAGCAGATTTGGGCTTTGGATGAGATAGCAGATAGAAATCTACTTTT
US-10-698-	550	GGGGAAAGAGTGGTGGCTTTGCTGCTAGAGAGAGTCTTTCTCAGGATCTTTTGTCTG
US-10-698-	550	GGGGAAAGAGTGGTGGCTTTGCTGCTAGAGAGAGTCTTTCTCAGGATCTTTTGTCTG
consensus		GGGGAAAGAGTGGTGGCTTTGCTGCTAGAGAGAGTCTTTCTCAGGATCTTTTGTCTG
US-10-698-	611	CTATATTTCTGGCTTAAAGAGAGAGTCTTTATGCGCAGGACTCACTTTTCCAAATGAACTCAT
US-10-698-	611	CTATATTTCTGGCTTAAAGAGAGAGTCTTTATGCGCAGGACTCACTTTTCCAAATGAACTCAT
consensus		CTATATTTCTGGCTTAAAGAGAGAGTCTTTATGCGCAGGACTCACTTTTCCAAATGAACTCAT
US-10-698-	672	CAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTGCTGATGTTCCAAATCTAGTAAAT
US-10-698-	672	CAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTGCTGATGTTCCAAATCTAGTAAAT
consensus		CAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTGCTGATGTTCCAAATCTAGTAAAT
US-10-698-	733	AAGCCTTTCAGAGAAAGGCTCAGGAGAGATCATTTGATGCTGCTCAAAATTTGAGCAGGAT
US-10-698-	733	AAGCCTTTCAGAGAAAGGCTCAGGAGAGATCATTTGATGCTGCTCAAAATTTGAGCAGGAT
consensus		AAGCCTTTCAGAGAAAGGCTCAGGAGAGATCATTTGATGCTGCTCAAAATTTGAGCAGGAT
US-10-698-	794	TTTTTAACAGAGCTTTGCCAGTTGGCTCATTTGGAATGAATTTGATGAAACAGTGA
US-10-698-	794	TTTTTAACAGAGCTTTGCCAGTTGGCTCATTTGGAATGAATTTGATGAAACAGTGA
consensus		TTTTTAACAGAGCTTTGCCAGTTGGCTCATTTGGAATGAATTTGATGAAACAGTGA
US-10-698-	855	CATTGAGTTTGTAGCTGACAGATTAATCTGTGGAACCTTGGATTTCTCAAGGTTTTTTCAGGCA
US-10-698-	855	CATTGAGTTTGTAGCTGACAGATTAATCTGTGGAACCTTGGATTTCTCAAGGTTTTTTCAGGCA
consensus		CATTGAGTTTGTAGCTGACAGATTAATCTGTGGAACCTTGGATTTCTCAAGGTTTTTTCAGGCA
US-10-698-	916	GAAAAATCCTTTTGAATTTTATGAAAAACATTTCTTTAGAGGAAAAACAAATTTCTTTGAGA
US-10-698-	916	GAAAAATCCTTTTGAATTTTATGAAAAACATTTCTTTAGAGGAAAAACAAATTTCTTTGAGA
consensus		GAAAAATCCTTTTGAATTTTATGAAAAACATTTCTTTAGAGGAAAAACAAATTTCTTTGAGA
US-10-698-	977	AACGAGTTTTCAGAGTATCAGCGTTTTTGGCAGTTATGCGCAGAAACCAAGATTAACGCTTTTCAC

```
US-10-698- 977 AACGAGTTTCAGAGTATCAGCGTTTTTGCAAGTTATGGCAGAAACACAGATAACGTCTTCAC
consensus   AACGAGTTTCAGAGTATCAGCGTTTTTGCAAGTTATGGCAGAAACACAGATAACGTCTTCAC

US-10-698- 1038 CTTGGATGCAGATTTT
              |||||
US-10-698- 1038 CTTGGATGCAGATTTT
consensus   CTTGGATGCAGATTTT
```

Alignment score = 1052.00
Scoring matrix:

	1	10
1	1052	
10		

IntelliGenetics

FastDB - Fast Pairwise Comparison of Sequences Release 5.4

Results file us-10-698-228-2.res made by bshears on Mon 26 Sep 105 11:00:28-PST.

Query sequence being compared:US-10-698-228-2 (1-1033)
 Number of sequences searched: 1
 Number of scores above cutoff: 1

Results of the initial comparison of US-10-698-228-2 (1-1053) with:
File : /home/bshears/swop*seq

SCORE	STDEV	Letter
117	0	N
234	0	U
351	0	M
468	0	B
584	0	E
701	0	R
818	0	O
935	0	F
1052	0	S
	5	E
	10	Q
	15	U
	20	E
	25	N
	30	C
	35	E
	40	S

PARAMETERS		
Similarity matrix	Unitary	K-tuple
Mismatch penalty	1	Joining penalty
Gap penalty	1.00	Window size
Gap size penalty	0.33	
Cutoff score	0	
Randomization group	0	

SEARCH STATISTICS		
Scores:	Mean 1052	Median 0
		Standard Deviation 0.00
Times:	CPU 00:00:00.00	Total Elapsed 00:00:00.00

Number of residues:	1053
Number of sequences searched:	1
Number of scores above cutoff:	1

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Score	Init.	Opt.
1. US-10-698-228-12	Sequence 12, Application	1053	1052	1052	0.00 0
1. US-10-698-228-2	(1-1053)				
US-10-698-228-12	Sequence 12, Application US/10698228				
Sequence 12, Application US/10698228					
GENERAL INFORMATION:					
APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA					
TITLE OF INVENTION: New Protein and its DNA					
FILE REFERENCE: 2619WOOP					
CURRENT APPLICATION NUMBER: US/10/698,228					
CURRENT FILING DATE: 2003-10-30					
PRIORITY APPLICATION NUMBER: US/10/019,733					
PRIORITY FILING DATE: 2001-12-28					
PRIORITY APPLICATION NUMBER: JP 11-181131					
PRIORITY FILING DATE: 1999-06-28					
PRIORITY APPLICATION NUMBER: JP 11-192391					
PRIORITY FILING DATE: 1999-07-06					
PRIORITY APPLICATION NUMBER: JP 2000-017770					
PRIORITY FILING DATE: 2000-01-21					
NUMBER OF SEQ ID NOS: 14					
SEQ ID NO 12					
LENGTH: 1053					
TYPE: DNA					
ORGANISM: Homo sapiens					
FEATURE:					
OTHER INFORMATION:					

TACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATTGAAACCATGCCCTATGTATTAGAAAAAA
|||||
|||||
TACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATTGAAACCATGCCCTATGTATTAGAAAAAA
440 450 460 470 480 490 500
510 520 530 540 550 560 570
GCAGATTGGGCTTTGCCATGGATAGCAGATAGAAAAATCTACTTTTGGGAAAGAGTGGTGGCCTTTGCTGCT
|||||
GCAGATTGGGCTTTGCCATGGATAGCAGATAGAAAAATCTACTTTTGGGAAAGAGTGGTGGCCTTTGCTGCT
510 520 530 540 550 560 570
580 590 600 610 620 630 640
GTAGAAGGAGTTTCTTCTCAGGATCTTTTGGCTCTATATTCTGGCTAAAGAGAGAGTCTTATGCCAGGA
|||||
GTAGAAGGAGTTTCTTCTCAGGATCTTTTGGCTCTATATTCTGGCTAAAGAGAGAGTCTTATGCCAGGA
580 590 600 610 620 630 640
650 660 670 680 690 700 710 720
CTCACTTTTCCAAATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGGCTTGCCTGATGTTCCAA
|||||
CTCACTTTTCCAAATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGGCTTGCCTGATGTTCCAA
650 660 670 680 690 700 710 720
730 740 750 760 770 780 790
TACTTAGTAATAAGCCTTCAGAGAAAGGCTCAGGGAGATCAATTGTTGACTGTCTCAAAAATTCAGCAGGAG
|||||
TACTTAGTAATAAGCCTTCAGAGAAAGGCTCAGGGAGATCAATTGTTGACTGTCTCAAAAATTCAGCAGGAG
730 740 750 760 770 780 790
800 810 820 830 840 850 860
TTTTTAACAGAGCCTTGGCCAGTTGGCCTCAATTGGAATGAATTGCATTTTGCATGAAACAGTACATTCAGTTT
|||||
TTTTTAACAGAGCCTTGGCCAGTTGGCCTCAATTGGAATGAATTGCATTTTGCATGAAACAGTACATTCAGTTT
800 810 820 830 840 850 860
870 880 890 900 910 920 930
GTAGCTCAGAGATTACTTGTGGAACCTTGGATTCTCAAGAGTTTTCAGGGCAGAAAAATCCTTTTGATTTTATG
|||||
GTAGCTCAGAGATTACTTGTGGAACCTTGGATTCTCAAGAGTTTTCAGGGCAGAAAAATCCTTTTGATTTTATG
870 880 890 900 910 920 930
940 950 960 970 980 990 1000
GAAAAATTTCTTTAGAGGAAAAACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGGTT
|||||
GAAAAATTTCTTTAGAGGAAAAACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGGTT
940 950 960 970 980 990 1000
1010 1020 1030 1040 1050 X
ATGCCAGAAACACAGATACGCTTTCACCTTGGATGCAGATTTT
|||||
ATGCCAGAAACACAGATACGCTTTCACCTTGGATGCAGATTTT
1010 1020 1030 1040 1050 X

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2005, 21:42:15 ; Search time 3258 Seconds
(without alignments)
15660.940 Million cell updates/sec

Title: US-10-698-228-12
Perfect score: 1053
Sequence: 1 atggcgacccggaaggcc.....tcacctggatgcagatttt 1053

Scoring table: IDENTITY_NUC
Gapo 10.0, Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: GenEmbl.*

- 1: gb_ba.*
- 2: gb_hhg.*
- 3: gb_in.*
- 4: gb_on.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pt.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1053	100.0	1053	6	AR454877 Sequence
2	1053	100.0	1053	6	BD064773 Novel pro
3	1053	100.0	1053	6	BD093086 Novel pro
4	1053	100.0	1081	6	AR454870 Sequence
5	1053	100.0	1081	6	BD064766 Novel pro
6	1053	100.0	1081	6	BD093079 Novel pro
7	1051.4	99.8	1053	6	AR454868 Sequence
8	1051.4	99.8	1053	6	BD064764 Novel pro
9	1051.4	99.8	1053	6	BD093077 Novel pro
10	1051.4	99.8	1056	6	CQ714252 Sequence
11	1051.4	99.8	1056	9	AB166671 Homo sapi
12	1051.4	99.8	4955	6	AR454869 Sequence
13	1051.4	99.8	4955	6	BD064765 Novel pro
14	1051.4	99.8	4955	6	BD093078 Novel pro
15	1051.4	99.8	4955	9	AB036063 Homo sapi
16	1048.2	99.5	1601	6	BD156916 Primer fo
17	1048.2	99.5	1601	6	AX877905 Sequence
18	1048.2	99.5	1601	9	AK001965 Homo sapi
19	896.8	85.2	4532	10	BC058103 Mus muscu

20	853.4	81.0	4571	9	BC042468 Homo sapi
21	851.8	80.9	900	9	AB163437 Homo sapi
22	839.2	79.7	171737	10	AC122379 Mus muscu
23	580.6	55.1	2500	6	AR300103 Sequence
24	580.6	55.1	2500	6	AR438503 Sequence
25	580.6	55.1	2500	6	AR562011 Sequence
26	580	55.1	1794	9	S40301 Homo sapien
27	580	55.1	2500	6	AR380825 Sequence
28	580	55.1	2500	6	AX333242 Sequence
29	580	55.1	2500	6	AX333687 Sequence
30	580	55.1	2500	6	AX334345 Sequence
31	580	55.1	2500	9	HSRR2SS X59618 H.sapiens R
32	578.4	54.9	1649	9	BC001886 Homo sapi
33	578.4	54.9	1653	9	BC030154 Homo sapi
34	578.4	54.9	3039	9	AK092671 Homo sapi
35	578.2	54.9	1562	5	BC061353 Xenopus t
36	573.6	54.5	217774	2	AK120603 Rattus no
37	573.2	54.4	3203	9	AK123010 Homo sapi
38	573	54.4	1538	5	CR762018 Xenopus t
39	573	54.4	1611	5	CR848103 Xenopus t
40	573	54.4	1647	5	BC080161 Xenopus t
41	570.4	54.2	2111	10	MUSRNRM2A M14223 Mouse ribon
42	568.8	54.0	1170	6	E14825 cDNA encodi
43	566.2	53.8	1570	5	BC072071 Xenopus l
44	564.2	53.6	2057	5	BC044355 Danio rer
45	562.6	53.4	1328	5	DRU57965 U57965 Danio rerio

ALIGNMENTS

RESULT 1
AR454877
LOCUS AR454877 1053 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 12 from patent US 6682917.
ACCESSION AR454877
VERSION AR454877.1 GI:42688832
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1053)
AUTHORS Nakamura,Y., Arakawa,H. and Tanaka,H.
TITILE Protein having a ribonucleotide Reductase activity and a DNA thereof
JOURNAL Patent: US 6682917-A 12 27-JAN-2004;
FEATURES Location/Qualifiers
1..1053
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN		Query Match	100.0%; Score 1053; DB 6; Length 1053;
		Best Local Similarity	100.0%; Pred. No. 6.4e-255;
		Matches 1053; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1	ATGGGCGACCCGGAAGCGCGGCTGGATCAGGATGAGATCATCTTCA	60
DB	1	ATGGGCGACCCGGAAGCGCGGCTGGATCAGGATGAGATCATCTTCA	60
QY	61	GACACCAAGCAAGTGAATTAAGTCAATGAAGGCCACTCTCTAAGAAAGAGTTCTCGC	120
DB	61	GACACCAAGCAAGTGAATTAAGTCAATGAAGGCCACTCTCTAAGAAAGAGTTCTCGC	120
QY	121	CGGTTTGTCATCTTTCCCAATCCAGTACCCCTGATATTGGAAATGTTAATACAGGCACAG	180
DB	121	CGGTTTGTCATCTTTCCCAATCCAGTACCCCTGATATTGGAAATGTTAATACAGGCACAG	180
QY	181	GCTTCCTCTCGACAGCAGAAGAGTTGACTTATCAAGAGGATCTCCCTCACTGGAACAAG	240
DB	181	GCTTCCTCTCGACAGCAGAAGAGTTGACTTATCAAGAGGATCTCCCTCACTGGAACAAG	240
QY	241	CTTAAGCAGATGAGAGTACTTCTCATCTCTCATCTTTAGCCCTTTTGGAGCCAGTGAT	300

Db	241	CTTAAAGCAGATGAGAGTACTTCTCATCTCACACTTAGCCCTTTTTCAGCCAGTGAT	300
Qy	301	GGAATTTGTAATGAAAAATTTGGTGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT	360
Db	301	GGAATTTGTAATGAAAAATTTGGTGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT	360
Qy	361	CGCTGTTCTATGGCTTTCAAAATCTCATCGAAGTGTTCACCTCAGAGATGTACAGTTTG	420
Db	361	CGCTGTTCTATGGCTTTCAAAATCTCATCGAAGTGTTCACCTCAGAGATGTACAGTTTG	420
Qy	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAAATTTTATTAATCAATTTGAA	480
Db	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAAATTTTATTAATCAATTTGAA	480
Qy	481	ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTCGATGATAGCAGATGAAAA	540
Db	481	ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTCGATGATAGCAGATGAAAA	540
Qy	541	TCTACTGTTTGGGAAAGAGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA	600
Db	541	TCTACTGTTTGGGAAAGAGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA	600
Qy	601	TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTCC	660
Db	601	TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTCC	660
Qy	661	AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTGTGATGTTCCAA	720
Db	661	AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTGTGATGTTCCAA	720
Qy	721	TACTTAGTAAATAAGCCTTCAGAAAAAGGTCAGGAGATCATTTGTTGATGCTGCATA	780
Db	721	TACTTAGTAAATAAGCCTTCAGAAAAAGGTCAGGAGATCATTTGTTGATGCTGCATA	780
Qy	781	ATTGAGCAGGAGTCTTAAACAGAACCTTTGCCAGTTGGCTCATTTGGAATGAATTCATT	840
Db	781	ATTGAGCAGGAGTCTTAAACAGAACCTTTGCCAGTTGGCTCATTTGGAATGAATTCATT	840
Qy	841	TGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAATTCGATTC	900
Db	841	TGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAATTCGATTC	900
Qy	901	AAGGTTTTTCAGGCGAAGATCCCTTTGATTTTATGGAACATTTCTTTAGAGGAAA	960
Db	901	AAGGTTTTTCAGGCGAAGATCCCTTTGATTTTATGGAACATTTCTTTAGAGGAAA	960
Qy	961	ACAAATTTCTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTTATGCGAGAAC	1020
Db	961	ACAAATTTCTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTTATGCGAGAAC	1020
Qy	1021	ACAGATAACGTTTCACTTGGATGAGATTTT 1053	
Db	1021	ACAGATAACGTTTCACTTGGATGAGATTTT 1053	

RESULT 2
BD064773
LOCUS
DEFINITION Novel protein and DNA thereof.
ACCESSION BD064773
VERSION BD064773.1 GI:22610376
KEYWORDS JP 2001269184-A/10.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1053)
REFERENCE
AUTHORS Nakamura, Y., Arakawa, H. and Tanaka, H.
TITLE Novel protein and DNA thereof
JOURNAL Patent: JP 2001269184-A 10 02-OCT-2001;
YUSUKE NAKAMURA, TAKEDA CHEMICAL INDUSTRIES LTD
COMMENT OS Homo sapiens (human)

PN	JP	2001269184-A/10	
PD	02-OCT-2001	JP 2000192401	
PI	27-JUN-2000	JP 2000192401	
PF	YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC		
C12N15/09	A61K31/711	A61K38/00, A61K37/02, A61K45/00, A61K48/00, PC	
A61P35/00			
PC	A61P43/00, C07K16/32, C07K16/40, C12N1/15, C12N1/19, C12N1/21	PC	
C12N5/10, C12N9/02			
PC	C12P21/02, C12P21/08, G01N33/15, G01N33/50, G01N33/53	PC	
C12R1:191			
PC	(C12N15/09, C12R1:19), (C12N9/02, C12R1:91), (C12P21/02, C12R1:91),		
PC	(C12P21/08, C12R1:91), (C12N15/00, A61K37/02, A61K37/48, C12N5/00,		
PC	(C12N15/00, C12R1:91), (C12N15/00, C12R1:19)		
CC	Novel protein and DNA thereof		
FI	Key	Location/Qualifiers	
FT	source	1..1053	
FT	source	/organism='Homo sapiens (human)'	
FEATURES	Location/Qualifiers		
source	1..1053		
	/organism='Homo sapiens'		
	/mol_type='genomic DNA'		
	/db_xref='taxon:9606'		
ORIGIN			
Query Match	100.0%;	Score 1053;	DB 6; Length 1053;
Best Local Similarity	100.0%;	Pred. No. 6.4e-255;	
Matches 1053;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Qy	1	ATGGGCGACCCGGAAGGCGGAGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA	60
Db	1	ATGGGCGACCCGGAAGGCGGAGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA	60
Qy	61	GCACCAAGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC	120
Db	61	GCACCAAGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC	120
Qy	121	CGGTTTGTTCATCTTTCCATCCAGTACCTGATTTTGGAAATGTATTAACAGGACAG	180
Db	121	CGGTTTGTTCATCTTTCCATCCAGTACCTGATTTTGGAAATGTATTAACAGGACAG	180
Qy	181	GCTTCTCTTGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAAACAG	240
Db	181	GCTTCTCTTGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAAACAG	240
Qy	241	CTTAAAGCAGATGAGAAAGTCTTCACTCTCACTTACATCTTAGCCCTTTTTCAGCAGGAT	300
Db	241	CTTAAAGCAGATGAGAAAGTCTTCACTCTCACTTACATCTTAGCCCTTTTTCAGCAGGAT	300
Qy	301	GGAATTTGTAATGAAAAATTTGGTGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT	360
Db	301	GGAATTTGTAATGAAAAATTTGGTGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT	360
Qy	361	CGCTGTTCTATGGCTTTTCAAAATTTCTCATCGAAGTGTTCACCTCAGAGATGTACAGTTTG	420
Db	361	CGCTGTTCTATGGCTTTTCAAAATTTCTCATCGAAGTGTTCACCTCAGAGATGTACAGTTTG	420
Qy	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAAATTTTATTAATCAATTTGAA	480
Db	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAAATTTTATTAATCAATTTGAA	480
Qy	481	ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTCGATGATAGCAGATGAAAA	540
Db	481	ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTCGATGATAGCAGATGAAAA	540
Qy	541	TCTACTGTTTGGGAAAGAGTGGTGGCTTTTGTGCTGTAGAGGAGTTTCTTCTCAGGA	600
Db	541	TCTACTGTTTGGGAAAGAGTGGTGGCTTTTGTGCTGTAGAGGAGTTTCTTCTCAGGA	600
Qy	601	TCTTTTGTGCTATATTTCTGGCTAAAGAGAGGCTTTTATGCGAGGACTCACTTTTCC	660
Db	601	TCTTTTGTGCTATATTTCTGGCTAAAGAGAGGCTTTTATGCGAGGACTCACTTTTCC	660
Qy	661	AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTGTGATGTTCCAA	720

VERSION	AR454870.1	GI:42688825	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 1081)		
TITLE	Nakamura,Y., Arakawa,H. and Tanaka,H.		
PROTEIN	Protein having a ribonucleotide Reductase activity and a DNA		
THEOREF	thereof		
PATENT	Patent: US 6682917-A 4 27-JAN-2004;		
JOURNAL	Location/Qualifiers		
FEATURES	1..1081		
SOURCE	/organism="unknown"		
	/mol_type="genomic DNA"		
ORIGIN			
Query Match	100.0%;	Score 1053;	DB 6; Length 1081;
Best Local Similarity	100.0%;	Pred. No. 6.4e-255;	
Matches 1053;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	ATGGGCGACCCGGAAGCCGGAAGCGCGCGGCTGGATCAGGATGAGAGATCATCTTCA	60
Db	20	ATGGGCGACCCGGAAGCCGGAAGCGCGCGGCTGGATCAGGATGAGAGATCATCTTCA	79
QY	61	GACACCAACGAAGTGAATAAAGTCAAAATGAAGAGCCACTCCTAAGAAAGAGTTCTCGC	120
Db	80	GACACCAACGAAGTGAATAAAGTCAAAATGAAGAGCCACTCCTAAGAAAGAGTTCTCGC	139
QY	121	CGTTTGTCTATCTTCCAAATCAGTACCCCTGATATTTGGAAATGTATAAAGAGGACAG	180
Db	140	CGTTTGTCTATCTTCCAAATCAGTACCCCTGATATTTGGAAATGTATAAAGAGGACAG	199
QY	181	GCTTCCTCTCGACAGCAGAGAGTTGACATATCAAAAGGATCTCCCTCACTGGACAG	240
Db	200	GCTTCCTCTCGACAGCAGAGAGTTGACATATCAAAAGGATCTCCCTCACTGGACAG	259
QY	241	CTTAAGCAGATGAGAAGTACTTCTATCTCTACATCTTAGCCTTTTGGACGACGATGAT	300
-Db	260	CTTAAGCAGATGAGAAGTACTTCTATCTCTACATCTTAGCCTTTTGGACGACGATGAT	319
QY	301	GGAAATGTAATGAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCGAGTTCCAGAGCT	360
Db	320	GGAAATGTAATGAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCGAGTTCCAGAGCT	379
QY	361	CGCTGTTCTATGCTTCAAAATCTCATCGAGATGTTCTACTCAGAGATGACAGTTTG	420
Db	380	CGCTGTTCTATGCTTCAAAATCTCATCGAGATGTTCTACTCAGAGATGACAGTTTG	439
QY	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA	480
Db	440	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA	499
QY	481	ACCATGCGCTATGTTAAGAAAAAGCAGATTGGCGCTTGCATGGATAGCAGATAGAAAA	540
Db	500	ACCATGCGCTATGTTAAGAAAAAGCAGATTGGCGCTTGCATGGATAGCAGATAGAAAA	559
QY	541	TCTACTTTTGGGAAAGAGTGGTGGCTTGTCTGTAGAGAGGTTTCTTCTCAGGA	600
Db	560	TCTACTTTTGGGAAAGAGTGGTGGCTTGTCTGTAGAGAGGTTTCTTCTCAGGA	619
QY	601	TCTTTTGTGCTATATCTTGGCTAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTTC	660
Db	620	TCTTTTGTGCTATATCTTGGCTAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTTC	679
QY	661	AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTCTGCTGATGTTCAA	720
Db	680	AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTCTGCTGATGTTCAA	739
QY	721	TACTTAGTAAATAGCCCTTCAGAGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAAA	780
Db	740	TACTTAGTAAATAGCCCTTCAGAGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAAA	799
QY	781	ATTGAGCAGGAGTTTAAACAGAGCCCTTGCCAGTTGGCCCTCATTTGGAAATGAAATTGCATT	840
Db	800	ATTGAGCAGGAGTTTAAACAGAGCCCTTGCCAGTTGGCCCTCATTTGGAAATGAAATTGCATT	859
QY	841	TTGATGAAACAGTACATTTAGCTTTAGCTGACAGATTTACTTTGTGGAACCTTGATTTCTCA	900
Db	860	TTGATGAAACAGTACATTTAGCTTTAGCTGACAGATTTACTTTGTGGAACCTTGATTTCTCA	919
QY	901	AAGGTTTTTTCAGGCAGAAAAATCCTTTTGTATTTTATGAAAAACATTTCTTTAGAGGAAAA	960
Db	920	AAGGTTTTTTCAGGCAGAAAAATCCTTTTGTATTTTATGAAAAACATTTCTTTAGAGGAAAA	979
QY	961	ACAAATTTTCTTTGAGAAAAAGAGTTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC	1020
Db	980	ACAAATTTTCTTTGAGAAAAAGAGTTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC	1039
QY	1021	ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053	
Db	1040	ACAGATAACGCTTTCACCTTGGATGAGATTTT 1072	
RESULT 5			
BD064766			
LOCUS	BD064766	1081 bp	DNA linear PAT 27-AUG-2002
DEFINITION	Novel protein and DNA thereof.		
ACCESSION	BD064766		
VERSION	BD064766.1	GI:22610369	
KEYWORDS	JP 2001269184-A/3.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 1081)		
JOURNAL	Nakamura,Y., Arakawa,H. and Tanaka,H.		
COMMENT	Novel protein and DNA thereof		
	Patent: JP 2001269184-A 3 02-OCT-2001;		
	YUSUKE NAKAMURA, TAKEDA CHEMICAL INDUSTRIES LTD		
	OS Homo sapiens (human)		
	PN JP 2001269184-A/3		
	PD 02-OCT-2001		
	PF 27-JUN-2000 JP 2000192401		
	PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC		
	C12N15/09, A61K31/711, A61K38/00, A61K45/00, A61K48/00, PC		
	A61P35/00,		
	PC A61P43/00, C07K16/32, C07K16/40, C12N1/15, C12N1/19, C12N1/21 PC		
	, C12N5/10, C12N9/02,		
	PC C12P21/02, C12P21/08, G01N33/15, G01N33/50, G01N33/53 // (C12N15/09,		
	PC C12R1:91),		
	PC (C12N15/09, C12R1:19), (C12N9/02, C12R1:91), (C12P21/02, C12R1:91),		
	PC (C12P21/08, C12R1:91), C12N15/00, A61K37/02, A61K37/48, C12N5/00,		
	PC (C12N15/00, C12R1:91), (C12N15/00, C12R1:19)		
	CC Novel protein and DNA thereof		
	CC Key		
	FT source		
	1..1081		
	Location/Qualifiers		
	FT		
	1..1081		
	/organism='Homo sapiens (human)'		
FEATURES			
source			
	1..1081		
	/organism='Homo sapiens'		
	/mol_type='genomic DNA'		
	/db_xref='taxon:9606'		
ORIGIN			
Query Match	100.0%;	Score 1053;	DB 6; Length 1081;
Best Local Similarity	100.0%;	Pred. No. 6.4e-255;	
Matches 1053;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	ATGGGCGACCCGGAAGCCGGAAGCGCGCGGCTGGATCAGGATGAGAGATCATCTTCA	60
Db	20	ATGGGCGACCCGGAAGCCGGAAGCGCGCGGCTGGATCAGGATGAGAGATCATCTTCA	79
QY	61	GACACCAACGAAGTGAATAAAGTCAAAATGAAGAGCCACTCCTAAGAAAGAGTTCTCGC	120
Db	80	GACACCAACGAAGTGAATAAAGTCAAAATGAAGAGCCACTCCTAAGAAAGAGTTCTCGC	139
QY	121	CGTTTGTCTATCTTCCAAATCAGTACCCCTGATATTTGGAAATGTATAAAGAGGACAG	180
Db	140	CGTTTGTCTATCTTCCAAATCAGTACCCCTGATATTTGGAAATGTATAAAGAGGACAG	199
QY	181	GCTTCCTCTCGACAGCAGAGAGTTGACATATCAAAAGGATCTCCCTCACTGGACAG	240
Db	200	GCTTCCTCTCGACAGCAGAGAGTTGACATATCAAAAGGATCTCCCTCACTGGACAG	259
QY	241	CTTAAGCAGATGAGAAGTACTTCTATCTCTACATCTTAGCCTTTTGGACGACGATGAT	300
-Db	260	CTTAAGCAGATGAGAAGTACTTCTATCTCTACATCTTAGCCTTTTGGACGACGATGAT	319
QY	301	GGAAATGTAATGAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCGAGTTCCAGAGCT	360
Db	320	GGAAATGTAATGAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCGAGTTCCAGAGCT	379
QY	361	CGCTGTTCTATGCTTCAAAATCTCATCGAGATGTTCTACTCAGAGATGACAGTTTG	420
Db	380	CGCTGTTCTATGCTTCAAAATCTCATCGAGATGTTCTACTCAGAGATGACAGTTTG	439
QY	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA	480
Db	440	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA	499
QY	481	ACCATGCGCTATGTTAAGAAAAAGCAGATTGGCGCTTGCATGGATAGCAGATAGAAAA	540
Db	500	ACCATGCGCTATGTTAAGAAAAAGCAGATTGGCGCTTGCATGGATAGCAGATAGAAAA	559
QY	541	TCTACTTTTGGGAAAGAGTGGTGGCTTGTCTGTAGAGAGGTTTCTTCTCAGGA	600
Db	560	TCTACTTTTGGGAAAGAGTGGTGGCTTGTCTGTAGAGAGGTTTCTTCTCAGGA	619
QY	601	TCTTTTGTGCTATATCTTGGCTAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTTC	660
Db	620	TCTTTTGTGCTATATCTTGGCTAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTTC	679
QY	661	AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTCTGCTGATGTTCAA	720
Db	680	AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTCTGCTGATGTTCAA	739
QY	721	TACTTAGTAAATAGCCCTTCAGAGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAAA	780
Db	740	TACTTAGTAAATAGCCCTTCAGAGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAAA	799
QY	781	ATTGAGCAGGAGTTTAAACAGAGCCCTTGCCAGTTGGCCCTCATTTGGAAATGAAATTGCATT	840

ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE	1 (bases 1 to 1081)
JOURNAL	Nakamura, Y., Arakawa, H. and Tanaka, H.
	Novel protein and its DNA
	Patent: WO 0100799-A 3 04-JAN-2001;
	TAKEEDA CHEMICAL INDUSTRIES LTD, YUSUKE NAKAMURA, HIROFUMI ARAKAWA,
	HIROSHI TANAKA
COMMENT	OS Homo sapiens (human)
	PN WO 0100799-A/3
	ED 04-JAN-2001
	PF 27-JUN-2000 WO 2000JP004189
	PR 28-JUN-1999 JP 99P 181131, 06-JUL-1999 JP 99P 192391 PR
	21-JAN-2000 JP 00P 017770
	PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
	C12N9/04, C12N15/53, C12N1/15, C12N1/19, C12N1/21, C12N5/10 PC
	, C12P21/02, A61K38/44,
	PC "A61K48/00, C07K16/40, G01N33/53, C12Q1/26
	CC
FEATURES	Key Location/Qualifiers.
source	1..1081
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
ORIGIN	
	Query Match 100.0%; Score 1053; DB 6; Length 1081;
	Best Local Similarity 100.0%; Pred. No. 6.4e-255;
	Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 ATGGGGGACCGCGAAAGCGCGGCTGGATCAGATCAGAGATCATCTTCA 60
Db	20 ATGGGGGACCGCGAAAGCGCGGCTGGATCAGATCAGAGATCATCTTCA 79
QY	61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGGCCACTCCTAAGAAAGAGTTCTCGC 120
Db	80 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGGCCACTCCTAAGAAAGAGTTCTCGC 139
QY	121 CGGTTTGTCATCTTTCCAAATCAGTATCCCTGATATTTGGAAAATGTATAACAGGCACAG 180
Db	140 CGGTTTGTCATCTTTCCAAATCAGTATCCCTGATATTTGGAAAATGTATAACAGGCACAG 199
QY	181 GCTTCCTCTCGACAGCAGCAGAGGTTGACTTATCAAGAGTCTCCCTCACTGGAACAAG 240
Db	200 GCTTCCTCTCGACAGCAGCAGAGGTTGACTTATCAAGAGTCTCCCTCACTGGAACAAG 259
QY	241 CTTAAGCAGATGAGAAGTACTTCATCTCTCACATCTTAGCCCTTTTTCAGCCAGTGAT 300
Db	260 CTTAAGCAGATGAGAAGTACTTCATCTCTCACATCTTAGCCCTTTTTCAGCCAGTGAT 319
QY	301 GGAAATGTAATGAAAATTTGGTGAGCGCCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
Db	320 GGAAATGTAATGAAAATTTGGTGAGCGCCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 379
QY	361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAAGTGTTCACCTACAGATGTACAGTTTG 420
Db	380 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAAGTGTTCACCTACAGATGTACAGTTTG 439
QY	421 CTGTAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA 480
Db	440 CTGTAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA 499
QY	481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGGCCCTTCGGATGGATAGCAGATAGAAA 540
Db	500 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGGCCCTTCGGATGGATAGCAGATAGAAA 559
QY	541 TCTACTTTTGGGGAAGAGTGCTGCTTTGCTGCTGTAGAAGGAGTTTCTTCTCAGGA 600
Db	560 TCTACTTTTGGGGAAGAGTGCTGCTTTGCTGCTGTAGAAGGAGTTTCTTCTCAGGA 619
QY	601 TCTTTTGTGCTATATTTGGCTCAAGAAAGAGAGGTTCTTATGCCAGAGCTCACTTTTTC 660

[illegible]

```
COMMENT OS Homo sapiens (human)
PN: JP 2001269184-A/1
PD 02-OCT-2001
PF 27-JUN-2000 JP 2000192401
PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
C12N15/09, A61K31/711, A61K38/00, A61K38/43, A61K45/00, A61K48/00, PC
A61P35/00,
PC A61P43/00, C07K16/32, C07K16/40, C12N1/15, C12N1/19, C12N1/21 PC
, C12N5/10, C12N9/02,
PC C12P21/02, C12P21/08, G01N33/15, G01N33/50, G01N33/53// (C12N15/09,
PC C12R1:91),
PC (C12N15/09, C12R1:19), (C12N9/02, C12R1:91), (C12P21/02, C12R1:91),
PC (C12P21/08, C12R1:91), C12N15/00, A61K37/02, A61K37/48, C12N5/00,
PC (C12N15/00, C12R1:91), (C12N15/00, C12R1:19)
CC Novel protein and DNA thereof
FH Key Location/Qualifiers
FT source 1..1053
FT /organism='Homo sapiens (human)'.
FEATURES
source
1..1053
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.8%; Score 1051.4; DB 6; Length 1053;
Best Local Similarity 99.9%; Pred. No. 1.6e-254;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
1 ATGGCGGACCCGGAAGCCGGAAGCGCGGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 ATGGCGGACCCGGAAGCCGGAAGCGCGGGCTGGATCAGGATGAGAGATCATCTTCA 60
61 GACACCAAGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAAGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
121 CGGTTGTGTCATCTTCCGATCAGTACCTGATATTTGGAAAATGTATAACAGGCACAG 180
Db 121 CGGTTGTGTCATCTTCCGATCAGTACCTGATATTTGGAAAATGTATAACAGGCACAG 180
181 GCTTCCTTCTGGACAGAGAGAGTTGACATTTATCAAGAGTCTCCCTCACTGGAACAG 240
Db 181 GCTTCCTTCTGGACAGAGAGAGTTGACATTTATCAAGAGTCTCCCTCACTGGAACAG 240
241 CTTAAGCAGATGAGAAGTACTTCACTCTCTCACATCTTAGCCTTTTTCGAGCCAGTGAT 300
Db 241 CTTAAGCAGATGAGAAGTACTTCACTCTCTCACATCTTAGCCTTTTTCGAGCCAGTGAT 300
301 GGAATTGAAATGAAATTTGGTGAGCGCTTTAGTCAGAGGTCGAGGTTCCAGAGGCT 360
Db 301 GGAATTGAAATGAAATTTGGTGAGCGCTTTAGTCAGAGGTCGAGGTTCCAGAGGCT 360
361 CGCTGTTTCTATGCTTTCAAAATCTCATCGAGATGTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGCTTTCAAAATCTCATCGAGATGTTCACTCAGAGATGTACAGTTTG 420
421 CTGATAGACATTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTGAA 480
Db 421 CTGATAGACATTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTGAA 480
481 ACCATGCCCTATGTTAGAAAAAGAGAGATGGGCTTGCATGGATAGCAGATAGAAAA 540
Db 481 ACCATGCCCTATGTTAGAAAAAGAGAGATGGGCTTGCATGGATAGCAGATAGAAAA 540
541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGAGAGAGGTTCTTCTTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGAGAGAGGTTCTTCTTCAGGA 600
601 TCTTTTGTCTATATTTCTGGCTAAAGAAAGAGAGTCTTATGCCAGGACTCACATTTTCC 660
Db 601 TCTTTTGTCTATATTTCTGGCTAAAGAAAGAGAGTCTTATGCCAGGACTCACATTTTCC 660
```

```
QY 661 AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTCCCTGATGTTCCAA 720
Db 661 AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTCCCTGATGTTCCAA 720
QY 721 TACTTAGTAAATAAGCCTTTTACAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 780
Db 721 TACTTAGTAAATAAGCCTTTTACAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 780
QY 781 ATTGAGCAGAGATTTTAAACAGAAAGCCTTGCCAGTTGGCCCTCATTTGAATGAATTCGATT 840
Db 781 ATTGAGCAGAGATTTTAAACAGAAAGCCTTGCCAGTTGGCCCTCATTTGAATGAATTCGATT 840
QY 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTGTGGAACCTTGGATCTTCA 900
Db 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTGTGGAACCTTGGATCTTCA 900
QY 901 AAGGTTTTTTCAGCAGAAATCCTTTTGTATTTTATGAAAACATTTCTTTAGAGGAAAA 960
Db 901 AAGGTTTTTTCAGCAGAAATCCTTTTGTATTTTATGAAAACATTTCTTTAGAGGAAAA 960
QY 961 ACAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGGTTTTCAGTTATGGCAGAAACC 1020
Db 961 ACAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGGTTTTCAGTTATGGCAGAAACC 1020
QY 1021 ACAGATAAGCTCTTCACTTGGATGACAGATTTT 1053
Db 1021 ACAGATAAGCTCTTCACTTGGATGACAGATTTT 1053
RESULT 9
BD093077 Novel protein and its DNA.
LOCUS BD093077
DEFINITION BD093077
ACCESSION BD093077
VERSION BD093077.1 GI:22638665
KEYWORDS WO 0100799-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1053)
AUTHORS Nakamura, Y., Arakawa, H. and Tanaka, H.
TITLE Novel protein and its DNA
JOURNAL Patent: WO 0100799-A 1 04-JAN-2001;
TAKEDA CHEMICAL INDUSTRIES LTD, YUSUKE NAKAMURA, HIROFUMI ARAKAWA,
HIROSHI TANAKA
COMMENT OS Homo sapiens (human)
PN WO 0100799-A/1
PD 04-JAN-2001
PF 27-JUN-2000 WO 2000JP004189
PR 28-JUN-1999 JP 99P 181131, 06-JUL-1999 JP 99P 192391 PR
21-JAN-2000 JP 00P 017770
PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
C12N9/04, C12N15/53, C12N1/15, C12N1/19, C12N1/21, C12N5/10 PC
, C12P21/02, A61K38/44,
PC A61K48/00, C07K16/40, G01N33/53, C12Q1/26
CC
FH Key Location/Qualifiers.
FEATURES
source
1..1053
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.8%; Score 1051.4; DB 6; Length 1053;
Best Local Similarity 99.9%; Pred. No. 1.6e-254;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGCGGACCCGGAAGCCGGAAGCGCGGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 ATGGCGGACCCGGAAGCCGGAAGCGCGGGCTGGATCAGGATGAGAGATCATCTTCA 60
61 GACACCAAGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAAGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
121 CGGTTGTGTCATCTTCCGATCAGTACCTGATATTTGGAAAATGTATAACAGGCACAG 180
Db 121 CGGTTGTGTCATCTTCCGATCAGTACCTGATATTTGGAAAATGTATAACAGGCACAG 180
181 GCTTCCTTCTGGACAGAGAGAGTTGACATTTATCAAGAGTCTCCCTCACTGGAACAG 240
Db 181 GCTTCCTTCTGGACAGAGAGAGTTGACATTTATCAAGAGTCTCCCTCACTGGAACAG 240
241 CTTAAGCAGATGAGAAGTACTTCACTCTCTCACATCTTAGCCTTTTTCGAGCCAGTGAT 300
Db 241 CTTAAGCAGATGAGAAGTACTTCACTCTCTCACATCTTAGCCTTTTTCGAGCCAGTGAT 300
301 GGAATTGAAATGAAATTTGGTGAGCGCTTTAGTCAGAGGTCGAGGTTCCAGAGGCT 360
Db 301 GGAATTGAAATGAAATTTGGTGAGCGCTTTAGTCAGAGGTCGAGGTTCCAGAGGCT 360
361 CGCTGTTTCTATGCTTTCAAAATCTCATCGAGATGTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGCTTTCAAAATCTCATCGAGATGTTCACTCAGAGATGTACAGTTTG 420
421 CTGATAGACATTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTGAA 480
Db 421 CTGATAGACATTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTGAA 480
481 ACCATGCCCTATGTTAGAAAAAGAGAGATGGGCTTGCATGGATAGCAGATAGAAAA 540
Db 481 ACCATGCCCTATGTTAGAAAAAGAGAGATGGGCTTGCATGGATAGCAGATAGAAAA 540
541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGAGAGAGGTTCTTCTTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGAGAGAGGTTCTTCTTCAGGA 600
601 TCTTTTGTCTATATTTCTGGCTAAAGAAAGAGAGTCTTATGCCAGGACTCACATTTTCC 660
Db 601 TCTTTTGTCTATATTTCTGGCTAAAGAAAGAGAGTCTTATGCCAGGACTCACATTTTCC 660
```


Qy	721	TACTTAGTAAATAAGCCCTTCAGAAAGAGGGTCAGGAGATCAATGTTGATGCTGCTCAA	780
Db	721	TACTTAGTAAATAAGCCCTTCAGAAAGAGGGTCAGGAGATCAATGTTGATGCTGCTCAA	780
Qy	781	ATTGAGCAGGAGTTTAAACAGAAAGCCTTGCAGTTGGCTCATTTGGAATGAATTCGATT	840
Db	781	ATTGAGCAGGAGTTTAAACAGAAAGCCTTGCAGTTGGCTCATTTGGAATGAATTCGATT	840
Qy	841	TTGATGAACAGTACATTTAGTTCAGCTGCACAGATTACTTGTGGAACCTTGGATTCTCA	900
Db	841	TTGATGAACAGTACATTTAGTTCAGCTGCACAGATTACTTGTGGAACCTTGGATTCTCA	900
Qy	901	AAGGTTTTTCAGGCGAAGAAATCTTTTGATTTTATGGAAGAAATTTCTTTAGAGGAAAA	960
Db	901	AAGGTTTTTCAGGCGAAGAAATCTTTTGATTTTATGGAAGAAATTTCTTTAGAGGAAAA	960
Qy	961	ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTATGGCAGAAACC	1020
Db	961	ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTATGGCAGAAACC	1020
Qy	1021	ACAGATAAGCTCTTCACCTTGGATGCAGATTTT	1053
Db	1021	ACAGATAAGCTCTTCACCTTGGATGCAGATTTT	1053
RESULT 11			
AB166671			
LOCUS			
DEFINITION			
	Homo sapiens p53R2 mRNA for p53-inducible ribonucleotide reductase		
	small subunit 2, complete cds.		
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
	Homo sapiens (human)		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE			
AUTHORS			
TITLE			
	Ugai,H. and Yokoyama,K.K.		
JOURNAL			
	Homo sapiens p53R2 mRNA for p53-inducible ribonucleotide reductase		
	small subunit 2, complete cds		
	Published Only in Database (2004)		
REFERENCE			
AUTHORS			
TITLE			
	Ugai,H. and Yokoyama,K.K.		
JOURNAL			
	Direct Submission		
	Submitted (05-MAR-2004) Hideyo Ugai, RIKEN BioResource Center, Gene		
	Engineering Division; 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074,		
	Japan (E-mail:ugai@rtc.riken.go.jp, Tel:81-29-836-3612,		
	Fax:81-29-836-9120)		
FEATURES			
source			
	1..1056		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="p53R2.C19"		
	/cell_line="Saos-2"		
	/cell_type="Osteogenic sarcoma"		
gene			
	1..1056		
	/gene="p53R2"		
CDS			
	1..1056		
	/gene="p53R2"		
	/codon_start=1		
	/product="p53-inducible ribonucleotide reductase small		
	subunit 2"		
	/protein_id="BAD12267.1"		
	/db_xref="GI:45259569"		
	/translation="MGDPPEAAGLDQDSSDTHSEIKSNEEPLLRKSSRRFVI		
	FPQYDPIKMYQAQAFWAEVDLSKDLPHWNKLUKADKEYFISHLAFPAASDGI		
	VNNELVRFQSVQVPCAFQFQILINPHVSMYSLDITVIRDPKREFLENAIE		
	TMPYVKKKADWALRDKRSTFGERVFAAEVGFSGFAAIFWLKKRGLMPGLT		
	FSNELISRDGLHCDACLMFQVLVNPSPSERVREILVDVKIBOEPFLTEALPVGLIG		
	MNCILMKQYIEFVADRLLVELGFSKVFQENPPDFMENISLEGKTNFPEKRVSEYQRF		
	AVNAETDNTDVTILDADF"		

Query Match	99.8%;	Score	1051.4;	DB 9;	Length	1056;			
Best Local Similarity	99.9%;	Pred. No.	1.6e-254;						
Matches	1052;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
Qy	1	ATGGCGGACCCGGAAGGCGGCGGCTGGATCAGGATCAGGATGAGAGATCATCTTCA	60						
Db	1	ATGGCGGACCCGGAAGGCGGCGGCTGGATCAGGATCAGGATGAGAGATCATCTTCA	60						
Qy	61	GACACCAACGAAAGTGAATAAAGTCAAAATGAAGACCCACTCTCTAAGAAAGAGTTCTCGC	120						
Db	61	GACACCAACGAAAGTGAATAAAGTCAAAATGAAGACCCACTCTCTAAGAAAGAGTTCTCGC	120						
Qy	121	CGGTTTGTCACTCTTCCAAATCCAGTACCCTGATATTTGGAAAATGTATAACAGGCACAG	180						
Db	121	CGGTTTGTCACTCTTCCAAATCCAGTACCCTGATATTTGGAAAATGTATAACAGGCACAG	180						
Qy	181	GCTTCTCTTCTGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAACAAG	240						
Db	181	GCTTCTCTTCTGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAACAAG	240						
Qy	241	CTTAAAGCAGATGAGAAGTACTTTCATCTCTCACTCTTAGCCCTTTTTCAGAGCAGTGTAT	300						
Db	241	CTTAAAGCAGATGAGAAGTACTTTCATCTCTCACTCTTAGCCCTTTTTCAGAGCAGTGTAT	300						
Qy	301	GGAATTGTAATGAAATTTGGTGGAGCCCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT	360						
Db	301	GGAATTGTAATGAAATTTGGTGGAGCCCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT	360						
Qy	361	GGCTGTTCTATGGCTTTCAAATTTCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG	420						
Db	361	GGCTGTTCTATGGCTTTCAAATTTCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG	420						
Qy	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA	480						
Db	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA	480						
Qy	481	ACCATGCGCTATGTTAAGAAAAAGCAGATTGGGCTTCGGATGGATAGCAGATAGAAAA	540						
Db	481	ACCATGCGCTATGTTAAGAAAAAGCAGATTGGGCTTCGGATGGATAGCAGATAGAAAA	540						
Qy	541	TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAAGGAGTTTCTTCTCAGGA	600						
Db	541	TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAAGGAGTTTCTTCTCAGGA	600						
Qy	601	TCTTTTGTGCTATATTTCTGGCTAAAGAGAGGTCCTTATGCGCAGAGCTCACTTTTCC	660						
Db	601	TCTTTTGTGCTATATTTCTGGCTAAAGAGAGGTCCTTATGCGCAGAGCTCACTTTTCC	660						
Qy	661	AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTCTGCTGATGTTCCAA	720						
Db	661	AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTCTGCTGATGTTCCAA	720						
Qy	721	TACTTAGTAAATAAGCCCTTCAGAAAGAGGTCAGGAGATCATTTGTGATGCTGCTCAAA	780						
Db	721	TACTTAGTAAATAAGCCCTTCAGAAAGAGGTCAGGAGATCATTTGTGATGCTGCTCAAA	780						
Qy	781	ATTGAGCAGGAGTTTAAACAGAACCTTGCAGTTGGCTCATTTGGAATGAATTCGATT	840						
Db	781	ATTGAGCAGGAGTTTAAACAGAACCTTGCAGTTGGCTCATTTGGAATGAATTCGATT	840						
Qy	841	TTGATGAACAGTACATTTAGTTCAGCTGCACAGATTACTTGTGGAACCTTGGATTCTCA	900						
Db	841	TTGATGAACAGTACATTTAGTTCAGCTGCACAGATTACTTGTGGAACCTTGGATTCTCA	900						
Qy	901	AAGGTTTTTCAGGCGAAGAAATCTTTTGATTTTATGGAAGAAATTTCTTTAGAGGAAAA	960						
Db	901	AAGGTTTTTCAGGCGAAGAAATCTTTTGATTTTATGGAAGAAATTTCTTTAGAGGAAAA	960						
Qy	961	ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTATGGCAGAAACC	1020						
Db	961	ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTATGGCAGAAACC	1020						

Db	845	TC	TTTTGCTGCTATATTTCTGGCTAAAGAGAGAGGCTTTATGCGAGAGCTCACTTTTTC	904			
Qy	661	AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACCTTTCCTTGCCTGATGTTCCAA	720				
Db	905	AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACCTTTCCTTGCCTGATGTTCCAA	964				
Qy	721	TACTTAGTAATTAAGCCTTCAGAAAGAGGCTCAGGAGATCATTTGTTGATGCTGTCAAA	780				
Db	965	TACTTAGTAATTAAGCCTTCAGAAAGAGGCTCAGGAGATCATTTGTTGATGCTGTCAAA	1024				
Qy	781	ATTGAGCAGGAGTTTTTAAACAGAGCCTTGCACAGTTTGGCCCTCATTGGATGAATTGCATT	840				
Db	1025	ATTGAGCAGGAGTTTTTAAACAGAGCCTTGCACAGTTTGGCCCTCATTGGATGAATTGCATT	1084				
Qy	841	TTGATGAAACAGTACATTGAGTTTGTAGCTGCAGAGATTACTTGTGGAACCTTGGATTTCTCA	900				
Db	1085	TTGATGAAACAGTACATTGAGTTTGTAGCTGCAGAGATTACTTGTGGAACCTTGGATTTCTCA	1144				
Qy	901	AAGGTTTTTTCAGGCAGAAAAATCCCTTTTGGATTTTATGGAACAATTTCTTTAGAAAGGAAAA	960				
Db	1145	AAGGTTTTTTCAGGCAGAAAAATCCCTTTTGGATTTTATGGAACAATTTCTTTAGAAAGGAAAA	1204				
Qy	961	ACAAATTTCTTCAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAGAAC	1020				
Db	1205	ACAAATTTCTTCAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAGAAC	1264				
Qy	1021	ACAGATAACGTCCTTTCACCTTTGGATGCAGATTTT	1053				
Db	1265	ACAGATAACGTCCTTTCACCTTTGGATGCAGATTTT	1297				
RESULT 13							
LOCUS	BD064765	4955 bp DNA linear PAT 27-AUG-2002					
DEFINITION	Novel protein and DNA thereof.						
ACCESSION	BD064765						
VERSION	BD064765.1	GI:22610368					
KEYWORDS	JP 2001269184-A/2.						
SOURCE	Homo sapiens (human)						
ORGANISM	Homo sapiens						
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.						
AUTHORS	1 (bases 1 to 4955)						
TITLE	Novel protein and DNA thereof						
JOURNAL	Patent: JP 2001269184-A 2 02-OCT-2001:						
COMMENT	YUSUKE NAKAMURA, TAKEDA CHEMICAL INDUSTRIES LTD						
FEATURES	OS	Homo sapiens (human)					
	PN	JP 2001269184-A/2					
	PD	02-OCT-2001					
	PF	27-JUN-2000 JP 2000192401					
	PI	YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC					
	C12N15/09, A61K31/711, A61K38/00, A61K38/43, A61K45/00, A61K48/00, PC						
	A61P35/00,						
	PC	A61P43/00, C07K16/32, C07K16/40, C12N1/15, C12N1/19, C12N1/21 PC					
	C12N5/10, C12N9/02,						
	PC	C12P21/02, C12P21/08, G01N33/15, G01N33/50, G01N33/53/(C12N15/09, PC					
FEATURES	PC	C12P1/91)					
	PC	(C12N15/09, C12R1/19), (C12N9/02, C12R1/91), (C12P21/02, C12R1/91),					
	PC	(C12P21/08, C12R1/91), C12N15/00, A61K37/02, A61K37/48, C12N5/00,					
	PC	(C12N15/00, C12R1/91), (C12N15/00, C12R1/19)					
	CC	Novel protein and DNA thereof					
	EH	Key Location/Qualifiers					
	FT	source 1..4955					
	FT	Location/Qualifiers					
	source	1..4955					
		/organism="Homo sapiens"					
ORIGIN		/mol_type="genomic DNA"					
		/db_xref="taxon:9606"					
Query Match	Score 1051.4; DB 6; Length 4955;						

Best Local Similarity 99.9%; Pred. No. 1.5e-254; Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
QY	1	ATGGCGACCGGAAAGCGCGGCGCGGCTGGATCAGATGAGATCATCTTCA	60
Db	245	ATGGCGACCGGAAAGCGCGGCGCGGCTGGATCAGATGAGATCATCTTCA	304
QY	61	GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTAAGAAAGAGTTCTCGC	120
Db	305	GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTAAGAAAGAGTTCTCGC	364
QY	121	CGTTTGTCACTTTCCAAATCCAGTACCCTGATATTTGGAAATGTATAAACAGCACAG	180
Db	365	CGTTTGTCACTTTCCAAATCCAGTACCCTGATATTTGGAAATGTATAAACAGCACAG	424
QY	181	GCTTCTCTTGACAGACGAGAGGTTGACTTATCAAGGATCTCCCTCACTGGAAACAAG	240
Db	425	GCTTCTCTTGACAGACGAGAGGTTGACTTATCAAGGATCTCCCTCACTGGAAACAAG	484
QY	241	CTTAAAGCAGATGAGAAGTACTTCACTCTCATCTTAGCCCTTTTTCGAGCCAGTGAT	300
Db	485	CTTAAAGCAGATGAGAAGTACTTCACTCTCATCTTAGCCCTTTTTCGAGCCAGTGAT	544
QY	301	GGAAATGTAATCAAAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCT	360
Db	545	GGAAATGTAATCAAAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCT	604
QY	361	CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAAGATGTTTCACTCAGAGATGTACAGTTG	420
Db	605	CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAAGATGTTTCACTCAGAGATGTACAGTTG	664
QY	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGAATGAA	480

Db

1265

ACAGATAACGTCTTCCACCTTGGATGCAGATTTT 1297

RESULT 14

BD093078

LOCUS

BD093078

DEFINITION

Novel protein and its DNA.

ACCESSION

BD093078

VERSION

BD093078.1 GI:22638666

KEYWORDS

WO 0100799-A/2.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS

Nakamura,Y., Arakawa,H. and Tanaka,H.

TITLE

Novel protein and its DNA

JOURNAL

Patent: WO 0100799-A 2 04-JAN-2001;
TAKEDA CHEMICAL INDUSTRIES LTD, YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA

COMMENT

OS Homo sapiens (human)
PN WO 0100799-A/2
PD 04-JAN-2001
PF 27-JUN-2000 WO 2000JP004189
PR 28-JUN-1999 JP 99P 181131,06-JUL-1999 JP 99P 192391 PR
PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
C12N9/04, C12N15/53, C12N1/15, C12N1/19, C12N1/21, C12N5/10 PC
, C12P21/02, A61K38/44,
PC A61K48/00, C07K16/40, G01N33/53, C12Q1/26
CC

PH Key Location/Qualifiers.

FEATURES

source

1..4955
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 99.8%; Score 1051.4; DB 6; Length 4955;
Best Local Similarity 99.9%; Pred. No. 1.5e-254;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY

1

ATGGCGACCGGAAAGCGCGGCGCGGCTGGATCAGATGAGATCATCTTCA

60

Db

245

ATGGCGACCGGAAAGCGCGGCGCGGCTGGATCAGATGAGATCATCTTCA

304

QY

61

GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC

120

Db

305

GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC

364

QY

121

CGTTTGTCACTTTCCAAATCCAGTACCCTGATATTTGGAAATGTATAAACAGCACAG

180

Db

365

CGTTTGTCACTTTCCAAATCCAGTACCCTGATATTTGGAAATGTATAAACAGCACAG

424

QY

181

GCTTCTCTTGACAGACGAGAGTGTGACTTATCAAGGATCTCCCTCACTGGAAACAAG

240

Db

425

GCTTCTCTTGACAGACGAGAGTGTGACTTATCAAGGATCTCCCTCACTGGAAACAAG

484

QY

241

CTTAAAGCAGATGAGAAGTACTTCACTCTCATCTTAGCCCTTTTTCGAGCCAGTGAT

300

Db

485

CTTAAAGCAGATGAGAAGTACTTCACTCTCATCTTAGCCCTTTTTCGAGCCAGTGAT

544

QY

301

GGAAATGTAATCAAAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCT

360

Db

545

GGAAATGTAATCAAAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCT

604

QY

361

CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAAGATGTTTCACTCAGAGATGTACAGTTG

420

Db

605

CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAAGATGTTTCACTCAGAGATGTACAGTTG

664

QY

421

CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGAATGAA

480


```
Db 905 |||||AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTCCAA 964
Qy 721 TACTTAGTAATAAGCCTTTCAAGAAAGGGTCAGGAGATCATTTGTTGATCGTGTCAA 780
Db 965 TACTTAGTAATAAGCCTTTCAAGAAAGGGTCAGGAGATCATTTGTTGATCGTGTCAA 1024
Qy 781 ATTGAGCAGGAGTTTTTAAACAGAAAGCCTTGCCAGTTGGCCTCATTTGGAATGAATTGCATT 840
Db 1025 ATTGAGCAGGAGTTTTTAAACAGAAAGCCTTGCCAGTTGGCCTCATTTGGAATGAATTGCATT 1084
Qy 841 TTGATGAACAGTACATTTGATGAGTTTGTAGCTGACAGATTACTTTGTGGAACCTTGGATTCTCA 900
Db 1085 TTGATGAACAGTACATTTGATGAGTTTGTAGCTGACAGATTACTTTGTGGAACCTTGGATTCTCA 1144
Qy 901 AAGGTTTTTTCAGGCAGAAATCCCTTTTGTATTTTATGGAAACATTTCTTTAGAAAGGAAA 960
Db 1145 AAGGTTTTTTCAGGCAGAAATCCCTTTTGTATTTTATGGAAACATTTCTTTAGAAAGGAAA 1204
Qy 961 ACAATTTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGAGTTATGGCAGAAACC 1020
Db 1205 ACAATTTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGAGTTATGGCAGAAACC 1264
Qy 1021 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1053
Db 1265 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1297
```

Search completed: October 30, 2005, 01:51:31
Job time : 3261 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2005, 21:38:30 ; Search time 440.5 Seconds
(without alignments)
14150.937 Million cell updates/sec

Title: US-10-698-228-12
Perfect score: 1053
Sequence: 1 atggcgaccggaaagcc.....tcaccttgatgcagatttt 1053

Scoring table: IDENTITY_NUC
Gapop 10_0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1053	100.0	1053	4	Aaf32447 Human rib
2	1053	100.0	1081	4	Aaf32440 Human rib
3	1051.4	99.8	1053	4	Aaf32438 Human rib
4	1051.4	99.8	2596	3	Aaa12411 CDNA enco
5	1051.4	99.8	4955	4	Aaf32439 Human rib
6	1051.4	99.8	4955	13	Adr24210 Breast ca
7	1048.2	99.5	1601	4	Aah14924 Human cDN
8	650.6	61.8	706	13	Adq57092 Novel can
9	580	55.1	2482	4	Aas44917 Human con
10	580	55.1	2500	6	Abi65414 Lung canc
11	580	55.1	2500	6	Abi66517 Lung canc
12	580	55.1	2500	6	Abi65859 Lung canc
13	580	55.1	2500	8	Abx10335 DNA enco
14	580	55.1	2500	11	Adi32044 Human cDN
15	580	55.1	2500	12	Adn04443 Antipsori
16	580	55.1	2500	12	Adq09273 Human RRM
17	580	55.1	2500	13	Adn37637 Tumour-as
18	580	55.1	3393	12	Adn03788 Antipsori
19	580	55.1	3393	12	Ado19225 Human PRO
20	578.4	54.9	1989	3	Aac78111 Human can

21	578.4	54.9	2215	12	ADK70302	Adk70302 Respirato
22	578.4	54.9	2216	10	ADJ56536	Adj56536 Murine cD
23	578.4	54.9	3294	4	Aah73225	Aah73225 Human cer
24	569	54.0	2641	4	AAS44745	Aas44745 Human ful
25	568.8	54.0	1170	2	AAV05641	Aav05641 Human rib
26	562.6	53.4	1328	12	ADO57308	Ado57308 DNA enco
27	547.8	52.0	978	4	AAS41006	Aas41006 CDNA enco
28	545.4	51.8	977	6	ABL90228	Abi90228 Human pol
29	527.2	50.1	608	4	Aah07707	Aah07707 Human cDN
30	514	48.8	1371	5	AAS79474	Aas79474 DNA enco
31	463	44.0	1289	4	ABLI14627	Abli14627 Drosophil
32	442	42.0	1146	6	AAS62387	Aas62387 CDNA sequ
33	439.4	41.7	481	2	AAx51874	Aax51874 Human sec
34	433.8	41.2	1218	8	ABT20895	Abt20895 Aspergill
35	429.8	40.8	14176	2	AAT84564	Aat84564 Swinepox
36	429.8	40.8	14176	4	AAF84949	Aaf84949 Reverse c
37	429.8	40.8	14176	4	AAF84948	Aaf84948 Nucleotid
38	420.4	39.9	3945	4	ABLI14626	Abli14626 Drosophil
39	404.8	38.4	963	13	ADS58378	Ads58378 Bacterial
40	398.6	37.9	1146	8	ABT19075	Abt19075 Aspergill
41	394.2	37.4	1173	13	ADT47814	Adt47814 Bacterial
42	393.6	37.4	1242	6	ABZ32250	Abz32250 Candida a
43	388.2	36.9	1292	13	ADS49379	Ads49379 Bacterial
44	382	36.3	1200	10	ACC61238	Acc61238 Gene sequ
45	382	36.3	1200	10	ADK63241	Adk63241 Disease t

ALIGNMENTS

RESULT 1
AAF32447
ID AAF32447 standard; cDNA; 1053 BP.
XX
AC AAF32447;
XX
DT 18-APR-2001 (first entry)
XX
DE Human ribonucleotide reductase related nucleotide sequence SEQ ID NO:12.
KW Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.
XX
OS Homo sapiens.
XX
PN WO200100799-A1.
XX
PD 04-JAN-2001.
XX
PF 27-JUN-2000; 2000WO-JP004189.
XX
PR 28-JUN-1999; 99JP-00181131.
PR 06-JUL-1999; 99JP-00192391.
PR 21-JAN-2000; 2000JP-00017770.

XX
XX (TAKE) TAKEDA CHEM IND LTD.
XX (NAKA/) NAKAMURA Y.

XX
XX Nakamura Y, Arakawa H, Tanaka H;
XX WPI; 2001-112446/12.

XX
XX Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
XX diagnosis, treatment and prevention of cancer.

XX
XX Claim 6; Page 96-97; 102pp; Japanese.

XX
XX The present invention describes a human ribonucleotide reductase
XX designated TP53R2H. TP53R2H has cytostatic activity. TP53R2H is part of
XX the DNA repair mechanism and its activity is induced by p53. It can be
XX used for the treatment, prevention and diagnosis of a wide range of
XX cancers. The present sequence represents a human ribonucleotide reductase
XX related sequence which is given in the exemplification of the present
XX invention

XX SQ Sequence 1053 BP; 314 A; 191 C; 246 G; 302 T; 0 U; 0 Other;

Query Match 100.0%; Score 1053; DB 4; Length 1053;
 Best Local Similarity 100.0%; Pred. No. 9e-290;
 Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGACCCGGAAGGCGGAGCGGCTGGATCAGGATCAGAGATCATCTTCA 60
 DB 1 ATGGCGACCCGGAAGGCGGAGCGGCTGGATCAGGATCAGAGATCATCTTCA 60

QY 61 GACACCAACGAAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
 DB 61 GACACCAACGAAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120

QY 121 CGTTTGTGTCATCTTCCAAATCAGTACCTGATATTTGGAAAAATGTATAACAGGCACAG 180
 DB 121 CGTTTGTGTCATCTTCCAAATCAGTACCTGATATTTGGAAAAATGTATAACAGGCACAG 180

QY 181 GCTTCCTTCTGACAGCAGAGGTTGACTTATCAAGAGATCTCCCTCACCTGGAACAG 240
 DB 181 GCTTCCTTCTGACAGCAGAGGTTGACTTATCAAGAGATCTCCCTCACCTGGAACAG 240

QY 241 CTTAAGCAGATGAGAAGTACTTCTCTCATCTTACATCTTTAGCCCTTTTTCAGCCAGTGAT 300
 DB 241 CTTAAGCAGATGAGAAGTACTTCTCTCATCTTACATCTTTAGCCCTTTTTCAGCCAGTGAT 300

QY 301 GGAATTGTAATGAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCGAGTTCAGAGGT 360
 DB 301 GGAATTGTAATGAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCGAGTTCAGAGGT 360

QY 361 CGCTGTTCTATGCGTTTCAAAATCTCATCGAATGTTTCACTCAGAGTGTACAGTTG 420
 DB 361 CGCTGTTCTATGCGTTTCAAAATCTCATCGAATGTTTCACTCAGAGTGTACAGTTG 420

QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
 DB 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480

QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTGGATGGATAGCAGATGAAAA 540
 DB 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTGGATGGATAGCAGATGAAAA 540

QY 541 TCTACTTTTGGGAAGAGTGTGCGCTTGTCTGTAGAGGAGTTTCTTCTCAGA 600
 DB 541 TCTACTTTTGGGAAGAGTGTGCGCTTGTCTGTAGAGGAGTTTCTTCTCAGA 600

QY 601 TCTTTTGTCTATATTTCTGGCTAAGAAAGAGAGGTCTTATGCCAGACTCACTTTTCC 660
 DB 601 TCTTTTGTCTATATTTCTGGCTAAGAAAGAGAGGTCTTATGCCAGACTCACTTTTCC 660

QY 661 AATGAACATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGTGTTCCAA 720
 DB 661 AATGAACATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGTGTTCCAA 720

QY 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGCTCAGGAGATCATTTGTGATGCTGCAAA 780
 DB 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGCTCAGGAGATCATTTGTGATGCTGCAAA 780

QY 781 ATTGAGCAGAGGTTTTTAACAGAAGCCTTGGCCTTCATTTGGAATGAATTGCATT 840
 DB 781 ATTGAGCAGAGGTTTTTAACAGAAGCCTTGGCCTTCATTTGGAATGAATTGCATT 840

QY 841 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
 DB 841 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900

QY 901 AAGGTTTTTTCAGGCGAATAATCCTTTTGAATTTTATGGAATAATTTCTTTAGAGGAAAA 960
 DB 901 AAGGTTTTTTCAGGCGAATAATCCTTTTGAATTTTATGGAATAATTTCTTTAGAGGAAAA 960

QY 961 ACAAAATTTCTTGGAAAAAGGTTTTCAGAGTATCAGGTTTTTTCAGTTATGCGAGAAC 1020
 DB 961 ACAAAATTTCTTGGAAAAAGGTTTTCAGAGTATCAGGTTTTTTCAGTTATGCGAGAAC 1020

DB 961 ACAAAATTTCTTGGAAAAAGGTTTTCAGAGTATCAGGTTTTTTCAGTTATGCGAGAAC 1020

QY 1021 ACAGATAACGTTCTTACCTTGGATGAGATTTT 1053
 DB 1021 ACAGATAACGTTCTTACCTTGGATGAGATTTT 1053

RESULT 2
 AAF32440
 ID AAF32440 standard; cDNA; 1081 BP.
 XX AAF32440;
 XX AC
 XX DT 18-APR-2001 (first entry)
 XX DE Human ribonucleotide reductase related nucleotide sequence SEQ ID NO:4.
 XX KW Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.
 XX OS Homo sapiens.
 XX PN WO200100799-A1.
 XX PD 04-JAN-2001.
 XX PF 27-JUN-2000; 2000WO-JP004189.
 XX PR 28-JUN-1999; 99JP-00181131.
 XX PR 06-JUL-1999; 99JP-00192391.
 XX PR 21-JAN-2000; 2000JP-00017770.
 XX (TAKE) TAKEDA CHEM IND LTD.
 XX (NAKA/) NAKAMURA Y.
 XX PI Nakamura Y, Arakawa H, Tanaka H;
 XX WPI; 2001-112446/12.
 XX Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
 PT diagnosis, treatment and prevention of cancer.
 XX Example 2; Page 91; 102pp; Japanese.
 XX The present invention describes a human ribonucleotide reductase
 CC designated TP53R2H. TP53R2H has cytosolic activity. TP53R2H is part of
 CC the DNA repair mechanism and its activity is induced by p53. It can be
 CC used for the treatment, prevention and diagnosis of a wide range of
 CC cancers. The present sequence represents a human ribonucleotide reductase
 CC related sequence which is used in an example from the present invention
 XX SQ Sequence 1081 BP; 322 A; 199 C; 252 G; 308 T; 0 U; 0 Other;

Query Match 100.0%; Score 1053; DB 4; Length 1081;
 Best Local Similarity 100.0%; Pred. No. 9.1e-290;
 Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGGACCCGGAAGGCGGAGCGGCTGGATCAGGATGAGATCATCTTCA 60
 DB 20 ATGGGGACCCGGAAGGCGGAGCGGCTGGATCAGGATGAGATCATCTTCA 79

QY 61 GACACCAACGAAAGTGAATAAGTCAATGAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
 DB 80 GACACCAACGAAAGTGAATAAGTCAATGAGAGCCACTCTTAAGAAAGAGTTCTCGC 139

QY 121 CGTTTGTGTCATCTTCCAAATCAGTACCTGATATTTGGAAAAATGTATAACAGGCACAG 180
 DB 140 CGTTTGTGTCATCTTCCAAATCAGTACCTGATATTTGGAAAAATGTATAACAGGCACAG 199

QY 181 GCTTCCTTCTGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 240
 DB 200 GCTTCCTTCTGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 259

QY 241 CTTAAGCAGATGAGAAGTACTTCTCATCTTCTAGCCCTTTTTCAGCCAGTGAT 300

Db 260 CTTTAAAGCAGATCAGAAAGTACTTCTCTCTCACAATCTTAGCCCTTTTTCAGCCAGTGAT 319
QY 301 GGAATTTGTAATAAGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 360
Db 320 GGAATTTGTAATAAGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 379
QY 361 CGCTGTTTCTATGGCTTTCAAAATTCATCGAAATGTTTCCACACAGAGATGTACAGTTTG 420
Db 380 CGCTGTTTCTATGGCTTTCAAAATTCATCGAAATGTTTCCACACAGAGATGTACAGTTTG 439
QY 421 CTGATAGACACTTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAA 480
Db 440 CTGATAGACACTTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAA 499
QY 481 ACCATGCCCTATGTTAAGAAAGAGAGATTTGGCCCTTGGCATGGATAGCAGATAGAAAA 540
Db 500 ACCATGCCCTATGTTAAGAAAGAGAGATTTGGCCCTTGGCATGGATAGCAGATAGAAAA 559
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAGAGTTCCTCTCAGGA 600
Db 560 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAGAGTTCCTCTCAGGA 619
QY 601 TCTTTTGTCTGCTATATTTCTGGCTTAAAGAAAGAGAGTTCCTATGCGAGACTCACTTTTTC 660
Db 620 TCTTTTGTCTGCTATATTTCTGGCTTAAAGAAAGAGAGTTCCTATGCGAGACTCACTTTTTC 679
QY 661 AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGGCTGTGTTCCAA 720
Db 680 AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGGCTGTGTTCCAA 739
QY 721 TACTTAGTAAATAAGCCCTTCAGAAAGAGGTGAGGAGATCATTTGTTGATGCTGTCAA 780
Db 740 TACTTAGTAAATAAGCCCTTCAGAAAGAGGTGAGGAGATCATTTGTTGATGCTGTCAA 799
QY 781 ATTGACGAGAGTTTAAACAGAGCCCTTGCCAGTTGGCTCATTGGAATGATTCGATT 840
Db 800 ATTGACGAGAGTTTAAACAGAGCCCTTGCCAGTTGGCTCATTGGAATGATTCGATT 859
QY 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGCAGAGATTAATTTGGAACCTTGGAATCTCA 900
Db 860 TTGATGAAACAGTACATTCAGTTTGTAGCTGCAGAGATTAATTTGGAACCTTGGAATCTCA 919
QY 901 AAGTTTTCAGCAGAAATTCCTTTTGTATTTATGGAACCAATTTCTTTAGAGGAAAA 960
Db 920 AAGTTTTCAGCAGAAATTCCTTTTGTATTTATGGAACCAATTTCTTTAGAGGAAAA 979
QY 961 ACAAAATTTCTTGAGAAACGAGTTTCAGAGATATCAGCGTTTTCAGTTATGGCAGAAACC 1020
Db 980 ACAAAATTTCTTGAGAAACGAGTTTCAGAGATATCAGCGTTTTCAGTTATGGCAGAAACC 1039
QY 1021 ACAGATAAGCTCTTCACTTGGATGCAGATTTT 1053
Db 1040 ACAGATAAGCTCTTCACTTGGATGCAGATTTT 1072

RESULT 3

ID AAF32438 standard; cdna; 1053 BP.

XX AAF32438;

XX AAF32438;

DT 18-APR-2001 (first entry)

DE Human ribonucleotide reductase TP53R2H nucleotide sequence SEQ ID NO:2.

XX Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.

OS Homo sapiens.

XX WO200100799-A1.

XX 04-JAN-2001.

XX 27-JUN-2000; 2000WO-JP004189.
PF 28-JUN-1999; 99JP-00181131.
PR 06-JUL-1999; 99JP-00192391.
PR 21-JAN-2000; 2000JP-00017770.
XX (TAKE) TAKEDA CHEM IND LTD.
PA (NAKA/) NAKAMURA Y.
XX Nakamura Y, Arakawa H, Tanaka H;
PI WPI: 2001-112446/12.
XX P-PSDB; AAB69050.
DR Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
PT diagnosis, treatment and prevention of cancer.
PS Claim 5; Fig 1-3; 102pp; Japanese.
XX The present sequence encodes a human ribonucleotide reductase designated
CC TP53R2H. TP53R2H has cytosolic activity. TP53R2H is part of the DNA
CC repair mechanism and its activity is induced by p53. It can be used for
CC the treatment, prevention and diagnosis of a wide range of cancers
XX
SQ Sequence 1053 BP; 314 A; 192 C; 246 G; 301 T; 0 U; 0 Other;
Query Match 99.8%; Score 1051.4; DB 4; Length 1053;
Best Local Similarity 99.9%; Pred. No. 2.6e-289;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGGGACCCGGAAGGCGCGGCTGGATCAGATGAGATCATCTTCA 60
Db 1 ATGGGGACCCGGAAGGCGCGGCTGGATCAGATGAGATCATCTTCA 60
QY 61 GACACCAAGAGTGAATTAAGTCAATGAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAAGAGTGAATTAAGTCAATGAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
QY 121 CGCTTTGTCTATCTTTTCAATCCAGTACCTCTGATATTTGGAATGTATAACAGGCACAG 180
Db 121 CGCTTTGTCTATCTTTTCAATCCAGTACCTCTGATATTTGGAATGTATAACAGGCACAG 180
QY 181 GCTTCTCTCTGAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 240
Db 181 GCTTCTCTCTGAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 240
QY 241 CTTAAGAGCAGATCAGAGTACTTCTCATCTCTAGCTTTTTCAGCCAGTAT 300
Db 241 CTTAAGAGCAGATCAGAGTACTTCTCATCTCTAGCTTTTTCAGCCAGTAT 300
QY 301 GGAATTTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGCT 360
Db 301 GGAATTTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGCT 360
QY 361 CGCTGTTTCTATGGCTTTCAAAATTCATCGAGATGTTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGGCTTTCAAAATTCATCGAGATGTTTCACTCAGAGATGTACAGTTTG 420
QY 421 CTGATAGACACTTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAA 480
Db 421 CTGATAGACACTTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAA 480
QY 481 ACCATGCCCTATGTTAAGAAAGAGAGTGGTGGCTTTGGCATGGATAGCAGATAGAAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAGAGAGTGGTGGCTTTGGCATGGATAGCAGATAGAAAA 540
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAAAGAGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAAAGAGTTTCTTCTCAGGA 600
QY 601 TCTTTTGTCTGCTATATTTCTGGCTTAAAGAAAGAGAGTTCCTATGCGAGACTCACTTTTTC 660

Qy	721	TACTTAGTAAATAAGCCCTTCAGAAAGAGGTCAGGAGATCATTTGTGTGATGCTGTCAAA	780
Db	776	TACTTAGTAAATAAGCCCTTCAGAAAGAGGTCAGGAGATCATTTGTGTGATGCTGTCAAA	835
Qy	781	ATTGAGCAGAGATTTTTAAACAGAGCCCTTGCCAGTTGGCCTCATTTGGAATGAATTGCATT	840
Db	836	ATTGAGCAGAGATTTTTAAACAGAGCCCTTGCCAGTTGGCCTCATTTGGAATGAATTGCATT	895
Qy	841	TTGATGAAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACTTGGATTTCTCA	900
Db	896	TTGATGAAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACTTGGATTTCTCA	955
Qy	901	AAGGTTTTTTCAGGCAGAAAAATCCTTTTGATTTTATGAAAAACATTTCTTTAGAAAGGAAAA	960
Db	956	AAGGTTTTTTCAGGCAGAAAAATCCTTTTGATTTTATGAAAAACATTTCTTTAGAAAGGAAAA	1015
Qy	961	ACAAATTTCTTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTTGCGATTATGGCAGAAACC	1020
Db	1016	ACAAATTTCTTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTTGCGATTATGGCAGAAACC	1075
Qy	1021	ACAGATAACGTCCTTCACCTTGGATGCAGATTTT	1053
Db	1076	ACAGATAACGTCCTTCACCTTGGATGCAGATTTT	1108

RESULT 5	
AAF32439	
ID	AAF32439 standard; cDNA; 4955 BP.
XX	
XX	
AC	AAF32439;
XX	
XX	
DT	18-APR-2001 (first entry)
XX	
XX	
DE	Human ribonucleotide reductase related nucleotide sequence SEQ ID NO:3.
XX	
XX	
KW	Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.

XX homo sapiens.
XX WO200100799-A1.
XX
XX 04-JAN-2001.
XX
XX 27-JUN-2000; 2000WO-JP004189.
XX
XX 28-JUN-1999; 99JP-00181131.
XX 06-JUL-1999; 99JP-00192391.
XX 21-JAN-2000; 2000JP-00017770.
XX
XX (TAKE) TAKEDA CHEM IND LTD.
XX (NAKA/) NAKAMURA Y.
XX
XX Nakamura Y, Arakawa H, Tanaka H;
XX WPI; 2001-112446/12.
XX
XX
XX Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
XX diagnosis, treatment and prevention of cancer.
XX Example 2; Page 87-90; 102pp; Japanese.
XX

CC The present invention describes a human ribonucleotide reductase
CC designated TP53R2H. TP53R2H has cytosolic activity. TP53R2H is part of
CC the DNA repair mechanism and its activity is induced by p53. It can be
CC used for the treatment, prevention and diagnosis of a wide range of
CC cancers. The present sequence represents a human ribonucleotide reductase
CC related sequence which is used in an example from the present invention
XX
SQ Sequence 4955 BP; 1526 A; 803 C; 999 G; 1627 T; 0 U; 0 Other;

Query Match 99.8%; Score 1051.4; DB 4; Length 4955;
Best Local Similarity 99.9%; Pred. No. 5.2e-289;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	ATGGCGACCCGGAAAGCGCGCGCGGCTGGATCAGGATGAGAGATCATCTTCA	60
Db	245	ATGGCGACCCGGAAAGCGCGCGCGGCTGGATCAGGATGAGAGATCATCTTCA	304
Qy	61	GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC	120
Db	305	GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC	364
Qy	121	CGGTTTGTCTATCTTTCCAAATCCAGTACCCCTGATATTTGGAAATATGTAATAACACGGCCACAG	180
Db	365	CGGTTTGTCTATCTTTCCAAATCCAGTACCCCTGATATTTGGAAATATGTAATAACACGGCCACAG	424
Qy	181	GCTTCTCTTCTGGACAGCAGAGAGGTTGACTTATCAAAGGATCTCCCTCACTCGAAACAAG	240
Db	425	GCTTCTCTTCTGGACAGCAGAGAGGTTGACTTATCAAAGGATCTCCCTCACTCGAAACAAG	484
Qy	241	CTTAAAGCAGATGAGAAAGTACTTCACTCTCTCACTCTTAGCCTTTTTTGAGCCAGTGAT	300
Db	485	CTTAAAGCAGATGAGAAAGTACTTCACTCTCTCACTCTTAGCCTTTTTTGAGCCAGTGAT	544
Qy	301	GGAAATCTGTAAATGAATAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT	360
Db	545	GGAAATCTGTAAATGAATAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT	604
Qy	361	CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG	420
Db	605	CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG	664
Qy	421	CTGATGACACTTACATCAGAGATCCCAAGAAAGGGAAATTTTATTTAATGCAATTTGAA	480
Db	665	CTGATGACACTTACATCAGAGATCCCAAGAAAGGGAAATTTTATTTAATGCAATTTGAA	724
Qy	481	ACCATGCCCTATCTTAAGAAAAAGCAGATTGGCCCTTGCATGCGATAGCAGATAGAAAA	540
Db	725	ACCATGCCCTATCTTAAGAAAAAGCAGATTGGCCCTTGCATGCGATAGCAGATAGAAAA	784
Qy	541	TCTACTTTTTGGGAAAAGAGTGGTGGCCCTTGCCTGTAGAAGAGAGTTTTCTCTCTCAGGA	600
Db	785	TCTACTTTTTGGGAAAAGAGTGGTGGCCCTTGCCTGTGTAGAGAGAGTTTTCTCTCTCAGGA	844
Qy	601	TCCTTTTGCTGCTATATTCCTGGCTTAAAGAGAGAGGCTTATGCGCAGGACTCACTTTTTCC	660
Db	845	TCCTTTTGCTGCTATATTCCTGGCTTAAAGAGAGAGGCTTATGCGCAGGACTCACTTTTTCC	904
Qy	661	AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA	720
Db	905	AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA	964
Qy	721	TACTTAGTAAATAAGCCTTCAGAAAGAAAGGCTCAGGGAGATCAATGTTGATGCTGTCAAA	780
Db	965	TACTTAGTAAATAAGCCTTCAGAAAGAAAGGCTCAGGGAGATCAATGTTGATGCTGTCAAA	1024
Qy	781	ATTGAGCAGGAGTTTTTAACAGAGGCTTGCAGTTTGGCCTCATTTGGAATGAATTCGATT	840
Db	1025	ATTGAGCAGGAGTTTTTAACAGAGGCTTGCAGTTTGGCCTCATTTGGAATGAATTCGATT	1084
Qy	841	TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTGTGGAACTTGGAATTCCTCA	900
Db	1085	TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTGTGGAACTTGGAATTCCTCA	1144
Qy	901	AAGGTTTTTTCAGCGCAAAATCCTTTTGTATTTTATGGAATAATCTTTTATAGAGGAAAA	960
Db	1145	AAGGTTTTTTCAGCGCAAAATCCTTTTGTATTTTATGGAATAATCTTTTATAGAGGAAAA	1204
Qy	961	ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTTTCAGTTATGGCAGAAACC	1020
Db	1205	ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTTTCAGTTATGGCAGAAACC	1264
Qy	1021	ACAGATAAAGCTCTTACCTTGGATGCAGATTTT	1053
Db	1265	ACAGATAAAGCTCTTACCTTGGATGCAGATTTT	1297

```

RESULT 6
ADR24210
ID ADR24210 standard; DNA; 4955 BP.
XX AC
XX ADR24210;
XX 21-OCT-2004 (first entry)
XX DE Breast cancer prognosis marker #71.
XX KW ds; breast cancer; prognosis; gene expression; diagnosis.
XX OS Homo sapiens.
XX PN WO2004065545-A2.
XX PD 05-AUG-2004.
XX PF 15-JAN-2004; 2004WO-US001100.
XX PR 15-JAN-2003; 2003US-00342887.
XX PA (ROSE-) ROSETTA INPHARMATICS LLC.
XX PE (NECA-) NETHERLANDS CANCER INST.
XX PI Van't Veer LJ, He Y;
XX DR WPI; 2004-593473/57.
XX CC Classifying a breast cancer patient according to prognosis comprises
XX PT determining the similarity between the level of expression of each of
XX PT five genes in a cell sample taken from patient, to control levels.
XX PS Disclosure; SEQ ID NO 71; 226bp; English.
XX CC The invention relates to a method of classifying a breast cancer patient
XX CC according to prognosis by determining the similarity between the level of
XX CC expression of each of five genes for which markers are listed in the
XX CC specification, in a cell sample taken from the breast cancer patient, to
XX CC control levels of expression for each respective five genes to obtain a
XX CC patient similarity value. The methods are useful for classifying a breast
XX CC cancer patient according to prognosis. Kits and computer program products
XX CC are useful for data analysis using the diagnostic, prognostic and
XX CC statistical methods of the invention. This sequence corresponds to a
XX CC marker used in the method of the invention.
XX SQ Sequence 4955 BP; 1526 A; 803 C; 999 G; 1627 T; 0 U; 0 Other;

Query Match 99.8%; Score 1051.4; DB 13; Length 4955;
Best Local Similarity 99.9%; Pred. No. 5.2e-289;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGGACCCGGAAGGCGGAGCGCGGGCTGGATCAGATGAGATCATCTTCA 60
DB 245 ATGGCGGACCCGGAAGGCGGAGCGCGGGCTGGATCAGATGAGATCATCTTCA 304
QY 61 GACACCAACGAAGTGAATAAAGTCAAATGAAGGCCACTCTTAAGAAAGAGTTCTCGC 120
DB 305 GACACCAACGAAGTGAATAAAGTCAAATGAAGGCCACTCTTAAGAAAGAGTTCTCGC 364
QY 121 CGGTTTGTCTTCCATCCAGTACCTGTATTTGGAAATATATAACAGGCACAG 180
DB 365 CGGTTTGTCTTCCATCCAGTACCTGTATTTGGAAATATATAACAGGCACAG 424
QY 181 GCTTCTCTTGACAGCAGAGAGTGTGACTTATCAAGAGATCTCCCTCAGTGAACAG 240
DB 425 GCTTCTCTTGACAGCAGAGAGTGTGACTTATCAAGAGATCTCCCTCAGTGAACAG 484
QY 241 CTTAAGCAGATGAGAAGTACTTCTCTCATCTTTAGCCCTTTTTCAGCCAGTGTAT 300
DB 485 CTTAAGCAGATGAGAAGTACTTCTCTCATCTTTAGCCCTTTTTCAGCCAGTGTAT 544

```

```

QY 301 GGAATTTGTAATGAAAAATTTGGTGAGGCGCTTTAGTCAGGAGGTGAGGTTCCAGAGGCT 360
DB 545 GGAATTTGTAATGAAAAATTTGGTGAGGCGCTTTAGTCAGGAGGTGAGGTTCCAGAGGCT 604
QY 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAAGTGTTCATCTAGAGATGACAGTTTG 420
DB 605 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAAGTGTTCATCTAGAGATGACAGTTTG 664
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAATCAATTGAA 480
DB 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAATCAATTGAA 724
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTCGGCTTCGGATGAGATAGCAGATAGAAA 540
DB 725 ACCATGCCCTATGTTAAGAAAAAGCAGATTCGGCTTCGGATGAGATAGCAGATAGAAA 784
QY 541 TCTACTTTTGGGAAAAGAGTGTGCTTGTGTAGAGAGGAGTTTCTTCTCAGGA 600
DB 785 TCTACTTTTGGGAAAAGAGTGTGCTTGTGTAGAGAGGAGTTTCTTCTCAGGA 844
QY 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGCTTATGCCAGGACTCACTTTTCC 660
DB 845 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGCTTATGCCAGGACTCACTTTTCC 904
QY 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGCCTTGCCTGATTTCCAA 720
DB 905 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGCCTTGCCTGATTTCCAA 964
QY 721 TACTTTAGTAAATAAGCCTTTCAGAGAAAGGGTCAGGAGATCATTTGATGCTGTCAA 780
DB 965 TACTTTAGTAAATAAGCCTTTCAGAGAAAGGGTCAGGAGATCATTTGATGCTGTCAA 1024
QY 781 ATTGAGCAGGAGTTTAAACAGAGCCCTTGCAGTTGGCTCATTTGGAATGATTCGATT 840
DB 1025 ATTGAGCAGGAGTTTAAACAGAGCCCTTGCAGTTGGCTCATTTGGAATGATTCGATT 1084
QY 841 TTGATGAACACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGAACTTGATTTCTCA 900
DB 1085 TTGATGAACACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGAACTTGATTTCTCA 1144
QY 901 AAGGTTTTTTCAGGCGAGAAAATCCTTTTGTATTTTATGAAAAACATTTCTTTAGAGGAAA 960
DB 1145 AAGGTTTTTTCAGGCGAGAAAATCCTTTTGTATTTTATGAAAAACATTTCTTTAGAGGAAA 1204
QY 961 ACAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC 1020
DB 1205 ACAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC 1264
QY 1021 ACAGATAACGCTCTTCCACCTTGGATGAGATTTT 1053
DB 1265 ACAGATAACGCTCTTCCACCTTGGATGAGATTTT 1297

RESULT 7
AAH14924
ID AAH14924 standard; cDNA; 1601 BP.
XX AC AAH14924;
XX 26-JUN-2001 (first entry)
XX DE Human cDNA sequence SEQ ID NO:12810.
XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX OS Homo sapiens.
XX PN EP1074617-A2.
XX PD 07-FEB-2001.
XX PF 28-JUL-2000; 2000EP-00116126.
XX

```

PR 29-JUL-1999; 99JP-00248036.
 PR 27-AUG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.
 XX (HELI-) HELIX RES INST.
 PA
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayaashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 DR WPI; 2001-318749/34.
 XX
 XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.
 XX
 XX Claim 8; SEQ ID NO 12810; 2537pp + Sequence Listing; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dr primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention
 XX
 SQ Sequence 1601 BP; 484 A; 290 C; 348 G; 479 T; 0 U; 0 Other;
 Query Match 99.5%; Score 1048.2; DB 4; Length 1601;
 Best Local Similarity 99.7%; Pred. No. 2.5e-288;
 Matches 1050; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAAATGTTCTACTCAGAGATGTACAGTTTG 420
 Db |||||
 QY 400 CGCTGCTTCTATGGCTTTCAAAATCTCATCGAGAAATGTTCTACTCAGAGATGTACAGTTTG 459
 Db |||||
 QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATGAA 480
 Db |||||
 QY 460 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATGAA 519
 Db |||||
 QY 481 ACCATGCCCTTATGTTAAGAAAAGCAGATTGGCCCTTGCATGGATAGCAGATAGAAAA 540
 Db |||||
 QY 520 ACCATGCCCTTATGTTAAGAAAAGCAGATTGGCCCTTGCATGGATAGCAGATAGAAAA 579
 Db |||||
 QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTGTCTGTAGAAAGAGTTTCTTCTCAGCA 600
 Db |||||
 QY 580 TCTACTTTTGGGAAAGAGTGGTGGCTTGTCTGTAGAAAGAGTTTCTTCTCAGCA 639
 Db |||||
 QY 601 TCTTTTGTCTCTATATTTCTGGCTTAAAGAAAGAGAGGTCTTATGCCAGGACTCACATTTTCC 660
 Db |||||
 QY 640 TCTTTTGTCTCTATATTTCTGGCTTAAAGAAAGAGAGGTCTTATGCCAGGACTCACATTTTCC 699
 Db |||||
 QY 661 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA 720
 Db |||||
 QY 700 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA 759
 Db |||||
 QY 721 TACTTAGTAAATAAGCCCTTCAGAAAGAAAGGTTCAGGAGATCATTTGTTGATGCTGTCAA 780
 Db |||||
 QY 760 TACTTAGTAAATAAGCCCTTCAGAAAGAAAGGTTCAGGAGATCATTTGTTGATGCTGTCAA 819
 Db |||||
 QY 781 ATTGAGCAGAGATTTTAAACAGAAAGCCCTTGCCTGCTTGCCTCATTTGGAATGAATTCATT 840
 Db |||||
 QY 820 ATTGAGCAGAGATTTTAAACAGAAAGCCCTTGCCTGCTTGCCTCATTTGGAATGAATTCATT 879
 Db |||||
 QY 841 TTGATGAACAGTACATTTGATGTTGTAGCTGACAGATTAATGTTGGAACATTTGGAATTCATCA 900
 Db |||||
 QY 880 TTGATGAACAGTACATTTGATGTTGTAGCTGACAGATTAATGTTGGAACATTTGGAATTCATCA 939
 Db |||||
 QY 901 AGGTTTTCAGCAGAGAAATCCTTTGATTTTATGAAAACATTTCTTTAGAGGAAAA 960
 Db |||||
 QY 940 AGGTTTTCAGCAGAGAAATCCTTTGATTTTATGAAAACATTTCTTTAGAGGAAAA 999
 Db |||||
 QY 961 ACAAAATTTCTTTCAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGAGTTATGGCAGAAACC 1020
 Db |||||
 QY 1000 ACAAAATTTCTTTCAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGAGTTATGGCAGAAACC 1059
 Db |||||
 QY 1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053
 Db |||||
 QY 1060 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1092
 Db |||||
 RESULT 8
 ADQ57092
 ID ADQ57092 standard; DNA; 706 BP.
 XX
 AC ADQ57092;
 XX
 XX 21-OCT-2004 (first entry)
 DT
 XX Novel canine microarray-related DNA sequence SeqID8394.
 DE
 XX canine microarray; drug screening; toxicity assay;
 KW environmental pollutant; cellular response; gene expression profile;
 KW toxic response; liver necrosis; fatty liver disease;
 KW protein adduct formation; hepatitis; dog; ds.
 XX
 OS Canis familiaris.
 XX
 PN WO2004063324-A2.
 XX
 PD 29-JUL-2004.
 XX
 XX 05-MAY-2003; 2003WO-US013853.
 XX
 XX 03-MAY-2002; 2002US-0377240P.
 PR

XX (GENE-) GENE LOGIC INC.
PA (PFIZ) PFIZER PROD INC.
XX
PI Diggins JC, Porter M, Wei T;
XX WPI; 2004-561890/54.
XX
PT New isolated nucleic acid molecule, useful for drug screening and
PT toxicity assays or for assessing the impact, including toxicity, of a
PT compound, pharmaceutical agent or environmental pollutant on a cell or
PT living organism.
XX
PS Claim 1; SEQ ID NO 8394; 41pp; English.
XX
XX This invention is related to a novel isolated canine nucleic acid
CC sequences and the construction of canine microarrays containing a
CC significant portion of the canine genome. The isolated canine nucleic
CC acid sequences of the invention may be useful for drug screening and
CC toxicity assays. The invention is therefore useful for assessing the
CC impact, including toxicity, of a compound, pharmaceutical agent or
CC environmental pollutant on a cell or living organism. The methods are
CC useful for detecting genes that are up- or down-regulated in canines in a
CC disease state. The sequences are useful as diagnostic agents or markers
CC to detect a cellular response in a sample individually or as part of a
CC gene expression profile. It is also useful as a target for agents that
CC modulate gene expression or activity. The database is useful for
CC producing electronic Northern blots that allow the user to determine the cell
CC type or tissue in which a given gene is expressed and to allow
CC determination of the abundance or expression level of a given gene in a
CC particular tissue or cell. The methods are useful for determining the
CC similarity of a toxic response to one or more individual compounds. The
CC methods are useful for predicting at least one toxic response or the
CC likelihood that a compound or test agent will induce various specific
CC pathologies such as those of the liver (liver necrosis, fatty liver
CC disease, protein adduct formation or hepatitis), those of the kidney,
CC heart, brain or testes, or other pathologies associated with at least one
CC of the toxins. The methods are also useful for predicting or elucidating
CC the potential cellular pathways influenced, induced or modulated by the
CC compound or test agent due to the similarity of the expression profile
CC compared to the profile induced by a known toxin. The present sequence is
CC that of a canine DNA sequence which was claimed for use during the
CC production of a canine microarray of the invention.
XX
SQ Sequence 706 BP; 208 A; 113 C; 158 G; 227 T; 0 U; 0 Other;

Query Match 61.8%; Score 650.6; DB 13; Length 706;
Best Local Similarity 95.2%; Pred. No. 5.5e-175;
Matches 671; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 349 GTTCCAGAGGCTCGCTGTTTCTATGGCTTTTCAAAATCTCATCGAGAAATGTCACCTCAGAG 408
DB 1 GTTCCAGAGGCTCGCTGTTTCTATGGCTTTTCAAAATCTCATCGAGAAATGTCACCTCAGAG 60
QY 409 ATGTACAGTTGCTGATGACACTTACATCAGAGATCCCAAGAAAGGAATTTTATTT 468
DB 61 ATGTACAGTTGCTGATGACACTTACATCAGAGATCCCAAGAAAGGAATTTTATTT 120
QY 469 AATGCAATTGAAACCATGCCCTATGTTAAGAAAAAGCAGATTGGCCCTTGGCATGGATA 528
DB 121 AATGCAATTGAAACCATGCCATATGTTAAGAAAAAGCAGATTGGCCCTTACATGGATA 180
QY 529 GCAGATAGAAATCTACTTTTGGGAAAGAGTGGTGGCCCTTGTGCTGTAGAAAGAGTT 588
DB 181 GAAGATAGAAATCTACTTTTGGGAAAGAGTGGTGGCCCTTGTGCTGTAGAAAGAGTT 240
QY 589 TTTCTTCAGGATCTTTGCTGCTATATTCGGCTTAAGAAAGAGGCTTTATGCGCAGA 648
DB 241 TTTCTTCAGGATCTTTGCTGCTATATTCGGCTTAAGAAAGAGGCTTTATGCGCTGGA 300
QY 649 CTCACCTTTTCCAAATGAACTCATCAGCAGAGATGAAGGACTTCACGTGACCTTGTCTGC 708
DB 301 CTCACCTTTTCCAAATGAACTCATCAGCAGAGATGAAGGCTTCACGTGACCTTGTCTGC 360

QY 709 CTGATGCTTCCAATACCTTAGTAAATAGCCCTTCAGAAAGAGGCTCAGGGAGATCATTTGTT 768
DB 361 CTGATGCTTCCAATACCTTAGTAAATAGCCCTTCAGAAAGAGGCTCAGGGAGATCATTTGTT 420
QY 769 GATGCTGTCAAAATTTGAGCAGGAGTGTAAACAGAAAGCCCTTGCAGTTGGCCCTCATTTGGA 828
DB 421 AATGCTGTGAAATTTGAGCAGGAGTGTAAACAGAAAGCCCTTGCAGTTGGCCCTCATTTGGA 480
QY 829 ATGAATTTGATTTTGTGATGAAACAGTACATTTGATGTTGTAGCTGACAGATTTACTTTGGAA 888
DB 481 ATGAATTTGTTTGTGATGAAACAGTATATTTAGTTGTAGCTGACAGATTTACTTTGGAA 540
QY 889 CTTGGATTCTCAAAAGTGTTCAGGCAGAGAAATCCCTTTTGTATTTATGAAAAACATTTCT 948
DB 541 CTTGGATTCTCAAAAGTGTTCAGGCAGAGAAATCCCTTTTGTATTTATGAAAAACATTTCT 600
QY 949 TTAGAAGGAAAAACAAATTTCTTTTGAGAAACGAGTTTCAGAGTATCAGCGCTTTTGCAAGTT 1008
DB 601 TTAGAAGGAAAAACAAATTTCTTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCGCGTT 660
QY 1009 ATGGCAGAAACCCACAGATTAACGTTCTTTCACCTTGGATGCGAGATTTT 1053
DB 661 ATGGCAGAAACCCACAGATTAACGTTCTTTCACCTTGGATGCGAGATTTT 705

RESULT 9
AAS44917/c
ID AAS44917 standard; DNA; 2482 BP.
XX
AC AAS44917;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human contig polynucleotide sequence #170.
XX
KW Mammal; human; rhesus monkey; baker's yeast; fission yeast; Norway rat;
KW mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;
KW cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;
KW nervous system disorder; inflammatory disorder; cell differentiation; ds;
KW angiogenesis; stem cell growth factor; activin; inhibitor; cartilage; burn;
KW genetic disorder; bone regeneration; tendon; ligament; tissue repair;
KW cystostatic; antirheumatic; antiarthritic; vulnery; antiinflammatory;
KW antibacterial; immunosuppressive; vasotropic; antiparkinsonian;
KW neuroprotective; osteopathic; antidiabetic; antiasthmatic; antiallergic;
KW immunostimulant; analgesic; gene therapy.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200164834-A2.
XX
PD 07-SEP-2001.
XX
XX 26-FEB-2001; 2001WO-US004926.
PF
XX 28-FEB-2000; 2000US-00515126.
XX
PR 18-MAY-2000; 2000US-00577409.
PR
PR 17-JUN-2000; 2000US-00597707.
PR
PR 14-JUL-2000; 2000US-00616807.
PR
PR 19-SEP-2000; 2000US-00664641.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;
PI Drmanac R;
XX
XX WPI; 2001-589862/66.
DR
DR P-PSDB; AAU28017.
XX
XX Novel polypeptides and nucleic acids obtained from cDNA libraries
PT prepared from various human tissues, for diagnosis, treatment of cancer,
PT

PT neurological, inflammatory disorders and for use in arrays for detection.
XX
PS Claim 1; SEQ ID NO 514; 153bp; English.
XX

CC Sequences AAS44576-AAS44919 represent full-length polynucleotides and
CC contig polynucleotides encoding polypeptides of the invention. The DNA
CC and protein sequences are useful for the treatment, diagnosis and
CC prevention of various types of disorder in a mammalian subject such as a
CC human, dog, monkey, mouse, hamster or rat. The disorders include cancers
CC such as leukaemia, lymphoma and neuroblastoma, autoimmune disorders such
CC as multiple sclerosis, connective tissue disease, rheumatoid arthritis,
CC diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system
CC disorders such as Parkinson's disease, Alzheimer's disease, Huntington's
CC chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and
CC Wernicke disease, inflammatory disorders such as nephritis, Crohn's
CC disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory
CC bowel disease. The sequences exhibit activity relating to angiogenesis,
CC cell proliferation, cell differentiation, stem cell growth factor,
CC activin or inhibin. Therefore, they can be used to manipulate stem cells
CC in culture to give rise to neuroepithelial cells that can be used to
CC augment or replace cells damaged by illness, accidental damage or genetic
CC disorders. The sequences may also be used for regeneration of bone,
CC cartilage, tendons and ligaments and in tissue repair and burn healing.
CC Note: Some sequences for this patent did not form part of the printed
CC specification, but were obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 2482 BP; 707 A; 598 C; 522 G; 655 T; 0 U; 0 Other;

Query Match 55.1%; Score 580; DB 4; Length 2482;
Best Local Similarity 75.1%; Pred. No. 1.4e-154;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

QY 89 ATGAAGAGGACCTCTAAGAAAGATTCTGCGCGGTTTGTCATCTTTCCAAATCCAGTACC 148
DB 2086 AGGATGAGCGGCTGTGAGAGAAACCCCGCGCGCTTTGTCTATCTTCCCATCCAGTACC 2027
QY 149 CTGATATTTGGAAATGTATAACAGCAGCAGGCTTCTCTGGACAGAGAGAGGTTG 208
DB 2026 ATGATATCTGGCAGATGTATAAGAGGAGAGGCTTCTTTGGACCGCGGAGAGGTTG 1967
QY 209 ACTTATCAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAAGTCTTCATCT 268
DB 1966 ACCTCTCCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGATATTTATAT 1907
QY 269 CTCACATCTAGCCTTTTTCAGCCAGTGATGGAATGTAAATGAAAAATTTGGTGGAGC 328
DB 1906 CCCATGCTTGGCTTCTTTGACGCAAGCGATGGCATAGTAAATGAAAACTTGGTGGAGC 1847
QY 329 GCTTTAGTCAGGAGTGAGGTTCCAGAGGCTCGCTGTTCTATGGCTTCAAAATCTCA 388
DB 1846 GATTTAGCCAAAGATTCAGATTACAGAAGCCCGCTGTTCTATGGCTTCCAAATGGCA 1787
QY 389 TCAGAAATGTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAGATCCCA 448
DB 1786 TGGAAACATACATCTGAAATGTATAGTCTTCTTATGACACTTACATAAAGATCCCA 1727
QY 449 AGAAAGGGAATTTTATTTAATGCAATTTGAAACCATGCCCTATGTTAAGAAAAAGCAG 508
DB 1726 AAGAAAGGGAATTTCTCTTCAATGCCATTTGAAACGATGCTTGTGTCAAGAGAGGAGCAG 1667
QY 509 ATTGGCCCTTGGATGGATAGCAGATAGAAATCTACTTTTGGGGAAGAGTGGGCT 568
DB 1666 ACTGGCCCTTGGCTGGATTGGGGAACAAAGAGGCTACCTATGGTGAACGCTTGTAGCCT 1607
QY 569 TTGCTGCTAGAAAGAGGTTTCTTCTCAGGATCTTTTGTCTGCTATATTCTGCTAAAGA 628
DB 1606 TTGCTGCTAGAAAGGATTTCTTTTTCGGTTCCTTTTCGGTTCGATATTCTGCTCAAGA 1547
QY 629 AGAGAGGTTTATGCGAGGACTCAGTTTTCATTTTCAATGAATCATCAGCAGAGATGAAGGAC 688
DB 1546 AACGAGGACTGATGCTGCGCTCACATTTTCTAAATGAATCTTATAGCAGAGATGAGGTT 1487

QY 689 TTCACTGTGACTTTGCTTGGCTGATGTTCCAAATAGTAAATAGGCTTTCAGAAAGAA 748
DB 1486 TACACTGTGATTTTGGCTTGGCTGATGTTCAACACCTGTTACACAAACATCGGAGGAGA 1427
QY 749 GGGTCAGGAGATCATTTGTTGATGCTGTCAAAATTTGAGCAGGAGTTTTTAACAGAGCCT 808
DB 1426 GAGTAAGAGAAATAATTTATCAATGCTGTTCCGATAGAACAGGAGTTCCTCACTGAGGCT 1367
QY 809 TGCAGTTGGCCTCATTTGGAATGAATTTGATGAAACAGTACATTTGATTTGTAG 868
DB 1366 TGCCTTGAAGACTCATTTGGGATGAATTTGCACTCTAATGAAGCAATACATTTGATTTGTGG 1307
QY 869 CTGACAGATTACTTGTGGAACCTTGATTTCTCAAAAGTTTTTTCAGGCAGAAATCTTTTG 928
DB 1306 CAGACAGCTTATGCTGGAACCTTGATTTTACCAAGTTTTTTCAGAGTAGAAGAACCATTTG 1247
QY 929 ATTTATGGAACATTTCTTTTGAAGGAAAAACAAATTTCTTTGAGAAACGAGTTTCAG 988
DB 1246 ACTTTATGGAGATATTTTCACTGGAAGGAAAGACTTAATCTTTTGAAGAGAGTAGGCG 1187
QY 989 AGTATCAGCGTTTTCAGTTATGCGCAAAACCAAGATGATGCTTTCACCTTTGGATGAG 1048
DB 1186 AGTATCAGAGATGGGAGTGTGTCAGTCAAGTCCACAGAGAAATTTCTTTACCTTTGGATGCTG 1127
QY 1049 ATTT 1052
DB 1126 ACTT 1123

RESULT 10

ABL65414

ID ABL65414 standard; DNA; 2500 BP.

XX ABL65414;

XX 15-MAY-2002 (first entry)

XX Lung cancer related gene sequence SEQ ID NO:3751.

DE Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
XX stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.

XX Homo sapiens.

XX WO2001194629-A2.

XX 13-DEC-2001.

XX 30-MAY-2001; 2001WO-US010838.

XX 05-JUN-2000; 2000US-0209473P.

XX 05-JUN-2000; 2000US-0209531P.

XX 18-SEP-2000; 2000US-0233133P.

XX 18-SEP-2000; 2000US-0233617P.

XX 20-SEP-2000; 2000US-0234009P.

XX 20-SEP-2000; 2000US-0234034P.

XX 20-SEP-2000; 2000US-0234052P.

XX 22-SEP-2000; 2000US-0234509P.

XX 22-SEP-2000; 2000US-0234567P.

XX 25-SEP-2000; 2000US-0234923P.

XX 25-SEP-2000; 2000US-0234924P.

XX 25-SEP-2000; 2000US-0235077P.

XX 25-SEP-2000; 2000US-0235082P.

XX 25-SEP-2000; 2000US-0235134P.

XX 26-SEP-2000; 2000US-0235637P.

XX 26-SEP-2000; 2000US-0235638P.

XX 27-SEP-2000; 2000US-0235711P.

XX 27-SEP-2000; 2000US-0235720P.

XX 27-SEP-2000; 2000US-0235840P.

XX 27-SEP-2000; 2000US-0235863P.

PR 28-SEP-2000; 2000US-0236028P.
 PR 28-SEP-2000; 2000US-0236032P.
 PR 28-SEP-2000; 2000US-0236033P.
 PR 28-SEP-2000; 2000US-0236034P.
 PR 28-SEP-2000; 2000US-0236109P.
 PR 28-SEP-2000; 2000US-0236111P.
 PR 29-SEP-2000; 2000US-0236842P.
 PR 29-SEP-2000; 2000US-0236891P.
 PR 02-OCT-2000; 2000US-0237172P.
 PR 02-OCT-2000; 2000US-0237173P.
 PR 02-OCT-2000; 2000US-0237278P.
 PR 02-OCT-2000; 2000US-0237294P.
 PR 02-OCT-2000; 2000US-0237295P.
 PR 02-OCT-2000; 2000US-0237316P.
 PR 03-OCT-2000; 2000US-0237425P.
 PR 03-OCT-2000; 2000US-0237598P.
 PR 03-OCT-2000; 2000US-0237604P.
 PR 03-OCT-2000; 2000US-0237606P.
 PR 03-OCT-2000; 2000US-0237608P.
 PR 01-NOV-2000; 2000US-0244867P.
 PR 01-NOV-2000; 2000US-0245084P.
 XX
 PA (AVAL-) AVALON PHARM.
 XX
 XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 PI
 XX WPI; 2002-188264/24.
 XX
 XX Screening for anti-neoplastic agent involves exposing cells to a chemical
 PT agent to be tested for anti-neoplastic activity, and determining a change
 PT in expression of a gene of a signature gene set.
 XX
 PS Claim 1; SEQ ID NO 3751; 44pp; English.
 XX
 CC The present invention describes a method (M1) for screening for an anti-
 CC neoplastic agent. The method involves exposing cells to a chemical agent
 CC to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL6164
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening an
 CC anti-neoplastic agent, and can be used for producing a product which is
 CC the data collected with respect to the anti-neoplastic agent as a result
 CC of M1, and the data is sufficient to convey the chemical structure and/or
 CC properties of the agent. M1 can be used in the treatment of cancer such
 CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
 CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
 CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
 CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms
 CC tumour
 XX
 SQ Sequence 2500 BP; 679 A; 522 C; 595 G; 704 T; 0 U; 0 Other;
 Query Match 55.1%; Score 580; DB 6; Length 2500;
 Best Local Similarity 75.1%; Pred. No. 1.4e-154;
 Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;
 QY 89 ATGAGAGGCACTCTCTAAGAAAGAGTTCTCGCGCGGTTCCTCAATCCAGTACC 148
 Db 397 AGGATGAGCGCGTCTGAGAGAAACCCCGCGCGTTCCTCAATCCAGTACC 456
 QY 149 CTGATATTGGAAATGTATAACAGGCACAGGCTTCCTTCGGACAGAGAAGAGTTG 208
 Db 457 ATGATATCTGGCAGATGATAGAAGGCAGAGGCTTCCTTTGGACCGCGAGGAGTTG 516
 QY 209 ACTTATCAAGGATCTCCCTGATCGAACAAGCTTAAGCAGATGAGAAGTACTTCATCT 268
 Db 517 ACCTCTCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTATAT 576
 QY 269 CTCACATCTTAGCCCTTTTTCGAGCAGGTGATGGAATGTAATGAAATTTGGTGAGC 328

Db 577 CCCATGTTCTGGCTTTCTTTTTCAGCAAGCGATGGCATAGTAAATGAAACTTGGTGAGC 636
 QY 329 GCTTTAGTCAGAGGTGCGAGGTTTCAGAGGCTCGCTGTTCTTATGCTTTCAATTTCTCA 388
 Db 637 GATTTAGCCCAAGAAGTTTCAGATTACAGAAAGCCCGCTGTTCTTATGCTTCCAAATTTGCCA 696
 QY 389 TCGAGAAATGTTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAGATCCCA 448
 Db 697 TGGAAAAACATACATCTCTGAAATGTATAGTCTTCTTATGACACTTACATAAAGATCCCA 756
 QY 449 AGAAAAGGGAATTTTATTTAATGCAATTTGAAACCAATGCCCTATGTTAAGAAAAAGCAG 508
 Db 757 AAGAAAAGGGAATTTCTCTTCAATGCCATTGAAACGATGCCCTGTGTCAAGAAAGGCGAG 816
 QY 509 ATTGGCCCTTGGGATGGATAGCAGATAGAAAATCTACTTTTGGGGAAGAGTGGTGGCCT 568
 Db 817 ACTGGGCCCTTGGCGTGGATTGGGGACAAAGAGGCTACCTATGGTGAACGTTGTAGCCT 876
 QY 569 TTGCTGCTGTAGAAGGAGTTTCTTCTCAGGATCTTTTCTGCTGCTATATTCTGGCTAAAGA 628
 Db 877 TTGCTGCTGTAGAAGGAGTTTCTTCTCAGGATCTTTTCTGCTGCTATATTCTGGCTAAAGA 936
 QY 629 AGAGAGGTCTTATGCCAGGACTCACATTTTCCAATGAACCTCATCAGCAGAGATGAAGGAC 688
 Db 937 AACGAGGACTGATGCGTGGCCTCACATTTTCTAATGAACTTATTAGCAGAGATGAGGTT 996
 QY 689 TTCACCTGTGACTTTGCTTGGCTGATGTTTCCAATGACTTACTAGTAATAAGCCTTCAGNAGAA 748
 Db 997 TACACTGTGATTTTGTCTGCTGATGTTTCCAACACCTGGTACACAAACCATCGAGGAGA 1056
 QY 749 GGGTCAGGAGAGATCATTTGTTGATGCTGTCAAAATTCAGCAGGAGTGTTCCTTAAACAGAACCT 808
 Db 1057 GAGTAAGAGAAATATATCAATGCTGTTCCGATAGAACAGGAGTTCCTCCTCCTGAGGCT 1116
 QY 809 TGCCAGTTGGCCTCATTTGGAATGAATTTGATGAAACAGTACATTCAGTTTGTAG 868
 Db 1117 TGCCGTGGAAGCTCATTTGGATGATTTGCACTTAATGAAGCAATACATTTGATTTGG 1176
 QY 869 CTGACAGATTACTTTGGAACCTTGAATCTCAAGGTTTTCAGGAGAGAAATTCCTTTTG 928
 Db 1177 CAGACAGACTTATGCTGGAACCTGGGTTTTCAGGAGGTTTTCAGAGTGAAGACCCCATTTG 1236
 QY 929 ATTTTATGAAAACATTTTCTTTAGAGGAAACAAATTTCTTTTGAGAAACGAGTTTTCAG 988
 Db 1237 ACTTATGAGAAATATTTCTGGAAGGAAAGACTAATCTTTTGAGAGAGATGAGCG 1296
 QY 989 AGTATCAGCGTTTTCAGTTCAGTTCAGGAGAAACCAAGATTAACGCTTCCTCCTTGGATCAG 1048
 Db 1297 AGTATCAGAGGATGGGAGTGTCAAGTCCAACAGAGAAATTCCTTTTACCTTGGATCTG 1356
 QY 1049 ATTT 1052
 Db 1357 ACTT 1360
 RESULT 11
 ABL66517
 ID ABL66517 standard; DNA; 2500 BP.
 XX
 AC ABL66517;
 XX
 DT 15-MAY-2002 (first entry)
 XX
 DE Lung cancer related gene sequence SEQ ID NO:4854.
 XX
 KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200194629-A2.

[illegible]

Db 1297 ACTATCAGAGGATGGAGTGAATGTCAGTCCACAGAGAAATCTTTTACCTTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 12
ABL65859
ID ABL65859 standard; DNA; 2500 BP.
XX AC
XX ABL65859;
DT 15-MAY-2002 (first entry)
XX DE
XX Lung cancer related gene sequence SEQ ID NO:4196.
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
XX WO200194629-A2.
XX
XX PD 13-DEC-2001.
XX PF
XX 30-MAY-2001; 2001WO-US010838.
XX
XX 05-JUN-2000; 2000US-0209473P.
XX PR 05-JUN-2000; 2000US-0209531P.
XX PR 18-SEP-2000; 2000US-0233133P.
XX PR 18-SEP-2000; 2000US-0233617P.
XX PR 20-SEP-2000; 2000US-0234009P.
XX PR 20-SEP-2000; 2000US-0234034P.
XX PR 20-SEP-2000; 2000US-0234052P.
XX PR 22-SEP-2000; 2000US-0234509P.
XX PR 22-SEP-2000; 2000US-0234567P.
XX PR 25-SEP-2000; 2000US-0234923P.
XX PR 25-SEP-2000; 2000US-0234924P.
XX PR 25-SEP-2000; 2000US-0235077P.
XX PR 25-SEP-2000; 2000US-0235082P.
XX PR 25-SEP-2000; 2000US-0235134P.
XX PR 25-SEP-2000; 2000US-0235280P.
XX PR 26-SEP-2000; 2000US-0235637P.
XX PR 26-SEP-2000; 2000US-0235638P.
XX PR 27-SEP-2000; 2000US-0235711P.
XX PR 27-SEP-2000; 2000US-0235720P.
XX PR 27-SEP-2000; 2000US-0235840P.
XX PR 27-SEP-2000; 2000US-0235863P.
XX PR 28-SEP-2000; 2000US-0236028P.
XX PR 28-SEP-2000; 2000US-0236032P.
XX PR 28-SEP-2000; 2000US-0236033P.
XX PR 28-SEP-2000; 2000US-0236034P.
XX PR 28-SEP-2000; 2000US-0236109P.
XX PR 29-SEP-2000; 2000US-0236111P.
XX PR 29-SEP-2000; 2000US-0236842P.
XX PR 29-SEP-2000; 2000US-0236891P.
XX PR 02-OCT-2000; 2000US-0237172P.
XX PR 02-OCT-2000; 2000US-0237173P.
XX PR 02-OCT-2000; 2000US-0237278P.
XX PR 02-OCT-2000; 2000US-0237294P.
XX PR 02-OCT-2000; 2000US-0237295P.
XX PR 02-OCT-2000; 2000US-0237316P.
XX PR 03-OCT-2000; 2000US-0237425P.
XX PR 03-OCT-2000; 2000US-0237598P.
XX PR 03-OCT-2000; 2000US-0237604P.
XX PR 03-OCT-2000; 2000US-0237606P.
XX PR 03-OCT-2000; 2000US-0237608P.
XX PR 01-NOV-2000; 2000US-0244867P.
XX PR 01-NOV-2000; 2000US-0245084P.

XX PA (AVAL-) AVALON PHARM.
XX PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
XX PI Soppet DR, Weaver Z;
XX DR WPI; 2002-188264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a chemical
XX PT agent to be tested for anti-neoplastic activity, and determining a change
XX PT in expression of a gene of a signature gene set.
XX PS Claim 1; SEQ ID NO 4196; 44pp; English.
XX
XX The present invention describes a method (M1) for screening for an anti-
XX CC neoplastic agent. The method involves exposing cells to a chemical agent
XX CC to be tested for anti-neoplastic activity, determining a change in
XX CC expression of at least one gene (I) of a signature gene set, where (I)
XX CC comprises a sequence (S) selected from 8447 sequences given in ABL61664
XX CC to ABL70110, or is at least 95% identical to (S), where a change in
XX CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
XX CC activity and can be used in gene therapy. M1 can be used for screening an
XX CC anti-neoplastic agent, and can be used for producing a product which is
XX CC the data collected with respect to the anti-neoplastic agent as a result
XX CC of M1, and the data is sufficient to convey the chemical structure and/or
XX CC properties of the agent. M1 can be used in the treatment of cancer such
XX CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
XX CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
XX CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
XX CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
XX CC tumour
XX
XX SQ Sequence 2500 BP; 679 A; 522 C; 595 G; 704 T; 0 U; 0 Other;
Query Match 55.1%; Score 580; DB 6; Length 2500;
Best Local Similarity 75.1%; Pred. No. 1.4e-154;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;
QY 89 ATGAAGAGGCCTCTTAAGAAAGAGTTCGCGCGTTTGTTCATCTTCCAAATCCAGTACC 148
Db 397 AGGATGAGCGCTGCTGAGAGAAAACCCCGCGCTTTGTTCATCTTCCCATCGAGTACC 456
QY 149 CTGATATTTGGAATGTATTAACAGGCACAGCTTCTTCGACAGCAGAGAGGTG 208
Db 457 ATGATATCTGGCAGATGTATAGAAGGAGAGGCTTCTTTGGACCGCAGGAGGTG 516
QY 209 ACTTATCAAAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAAGTACTTCACT 268
Db 517 ACCTCTCAAGGACATTGAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT 576
QY 269 CTCACATCTTAGCCTTTTGTGAGCCAGTGAATGGAATGAAATGAAAAATTTGGTGAGC 328
Db 577 CCCATGTTCTGGCTTCTTTGACAGCAAGCGATGGCATAGTAATAAATGAAAACTTGGTGAGC 636
QY 329 GCTTTAGTCAGAGGTGCGAGTTCAGAGCTCGCTGTTTCTATGGCTTCAAAATCTCA 388
Db 637 GATTTAGCCAAGAAGTTTCAGATTACAGAAGCCGCTGTTTCTATGGCTTCAAAATGCCA 696
QY 389 TCGAGAATGTTCACTCAGAGATGTACAGTTTGTGTGATAGACACTTACATCAGAGATCCCA 448
Db 697 TGGAAAAACATACATTCTGAAATGTATAGTCTTTTATTGACACTTACATAAAGATCCCA 756
QY 449 AGAAAAAGGGAATTTTATTTAATGCAATGAAACCATGCCCTATGTTTAAGAAAAAGCAG 508
Db 757 AAGAAAAGGGAATTTCTCTTCATGCCATTGAAACGATGCCCTTGTGTCAAGAAGAGGCAG 816
QY 509 ATTGGCCTTGGCATGGATAGCAGATAGAAAAATCTACTTTTGGGAAAGAGTGTGGCCT 568
Db 817 ACTGGGCCCTTGGCGTGGATTGGGGACAAGAGAGCTACCTATGTGTGAACGTTGTAGCCT 876
QY 569 TTGCTGCTGTAGAAGGAGTTTCTTCTCAGGATCTTTTGTGCTGTATATTCTGGCTAAAGA 628
Db 877 TTGCTGAGTGGAAAGCAATTTCTTTTCCGGTCTTTTGTGCTGATATTCTGGCTCAAGA 936

QY 629 AGAGAGGCTTATGCGAGGACTCCTTTTCCAAATGAACTCATCAGCAGATGAGGAC 688
 Db 937 AACGAGGACTGATGCTGCGCTCACAATTTCTAATGAATCTATTAGCAGATGAGGTT 996
 QY 689 TTCACCTGTGACTTGTGCTGCTGATGTTCCAAATCTAGTAAATAGCCCTTCAGAGAAA 748
 Db 997 TACACTGTGATTTGCTTGCCTGATGTTCCAAACCTGCTGACACAAACCATCGGAGGA 1056
 QY 749 GGGTCAGGAGATCATTTGATGCTGCTCAAAATTCAGCAGGAGTTTAAACAGAGCCT 808
 Db 1057 GAGTAAGAGAAAATAATTCAATGCTGTTCCGATAGAACAGGAGTTCTCTCACTGAGCCT 1116
 QY 809 TCCAGTGTGCTCCTATGGAATGAATGCAATTTGATGAAACAGTACATTTGATG 868
 Db 1117 TCCCTGTGAGCTCATTTGGATGGAATGCACTTAAATGAAGCAATACATTTGATGTTGG 1176
 QY 869 CTGACAGATTTACTGTGGAACCTTGATCTCAAGGTTTTCAGGCAGAAAATCCTTTTG 928
 Db 1177 CAGACAGACTTATGCTGNACTGGTGGTTTTCAGAGGTTTTCAGATGAGAACCATTTG 1236
 QY 929 ATTTATGGAACAATTTCTTTAGAGGAAAAACAATAATTTTGTGAGAACCGATTTTCA 988
 Db 1237 ACTTTATGAGAAATATTTCACTGGAAGGAAAGACTAACTTCTTTGAGAGAGAGTAGGCG 1296
 QY 989 AGTATCAGGCTTTTGCAGTTATGCGAGAAACACAGATTAACGTTTCACTTGGATGCG 1048
 Db 1297 AGTATCAGAGGATGGAGTGATGTCAAAGTCCAACAGAGAAATCTTTTACCTTGGATGCTG 1356
 QY 1049 ATTT 1052
 Db 1357 ACTT 1360

RESULT 13

ID ABX10335 standard; DNA; 2500 BP.
 XX
 AC ABX10335;
 XX
 DT 28-JAN-2003 (first entry)
 XX
 DE DNA encoding protein differentially regulated in prostate cancer #4.
 XX
 KW Prostate cancer; gene expression; differential regulation;
 KW molecular marker; drug target; cancer detection; cancer diagnosis;
 KW cancer staging; cancer grading; cancer assessing; cancer monitoring;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200281638-A2.
 XX
 PD 17-OCT-2002.
 XX
 PF 08-APR-2002; 2002WO-US010824.
 XX
 PR 06-APR-2001; 2001US-0281731P.
 PR 06-APR-2001; 2001US-0281732P.
 XX
 PA (ORIG-) ORIGENE TECHNOLOGIES INC.
 XX
 PI Sun Z, Jay G;
 XX
 DR WPI; 2003-058520/05.
 DR P-PSDB; ABU07433.
 XX
 PT Novel genes which are differentially regulated in prostate cancer, useful
 PT for diagnosing prostate cancer in prostate tissue sample and assessing
 PT therapeutic or preventive intervention in prostate cancer patients.
 XX
 PS Claim 1; Page 71-72; 416pp; English.

CC The invention describes genes (I) which are differentially regulated in
 CC prostate cancer. (I) Is useful for diagnosing a prostate cancer in a
 CC sample comprising prostate tissue, which involves determining the number
 CC of target genes which are differentially-regulated in the sample, where
 CC the number is indicative of the probability that the sample comprises
 CC prostate cancer. (I) Is useful for assessing a therapeutic or preventive
 CC intervention in a subject having a prostate cancer, which involves
 CC determining the expression levels in a sample comprising prostate tissue
 CC of target genes which are differentially-regulated in prostate cancer.
 CC Preferably, the expression levels of at least 10 genes are determined.
 CC (I) is also useful for identifying agents that modulate a biological
 CC activity of a polypeptide differentially-regulated in prostate cancer
 CC cells, which involves contacting a polypeptide differentially-regulated
 CC in prostate cancer cells with a test agent under conditions effective for
 CC the test agent to modulate a biological activity of the polypeptide, and
 CC determining whether the test agent modulates the biological activity. (I)
 CC is useful as molecular markers, as drug targets, and for detecting,
 CC diagnosing, staging, grading, assessing, monitoring, prognosticating,
 CC preventing or treating, determining predisposition to diseases and
 CC conditions especially relating to prostate cancer. (I) and its expression
 CC products are used in the diagnostic test to assay for presence of cancer
 CC e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in
 CC blood etc. (I) is useful for assessing cancer e.g., to determine the type
 CC of cancer, its stage of development, the nature of genetic defect, etc.
 CC The polypeptide encoded by (I) can be used as target for therapy or drug
 CC discovery. (I) can also be used for expressing the polypeptide and thus
 CC for searching specific binding partners of the polypeptide. (I) is useful
 CC in therapeutic applications to treat prostate cancer. The identification
 CC of specific genes, and groups of genes, expressed in pathways
 CC physiologically relevant to prostate cancer permits the definition of
 CC functional and disease pathways and the delineation of targets in these
 CC pathways which are useful in diagnostic, therapeutic, and clinical
 CC applications. This sequence encodes a protein differentially regulated in
 CC prostate cancer
 XX
 SQ Sequence 2500 BP; 679 A; 522 C; 595 G; 704 T; 0 U; 0 Other;

Query Match 55.1%; Score 580; DB 8; Length 2500;
 Best Local Similarity 75.1%; Pred. No. 1.4e-154;
 Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;
 QY 89 ATGAGAGGCACTCTTAAGAAAGATTCTCGCGGTTTGTTCATCTTTCCATCCAGTACC 148
 Db 397 AGATGAGCGCTGCTGAGAGAAAACCCCGCGCTTTGTTCATCTTCCCATCGAGTACC 456
 QY 149 CTGATATTTGGAAAATGTATAAACAGCAGCAGGCTTCTTCTGACAGCAGAGAGGTTG 208
 Db 457 ATGATATCTGGCAGATGTATAGAACGACAGAGGCTTCTTCTTGGACCGCGAGAGGTTG 516
 QY 209 ACTTATCAAAGGATCTCCCTCACTGGAAACAAGCTTAAAGCAGATGAGAAGTACTCATCT 268
 Db 517 ACCTCTCCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT 576
 QY 269 CTCACATCTTAGCTTTTTCAGCCAGTGTGGAATTTAAATGAAATTTGGTGGAGC 328
 Db 577 CCCATGTTCTGCGCTTTCTTTGCAAGCAGCGATGGCATAGTAAATGAAACTTGGTGGAGC 636
 QY 329 GCTTTAGTCAGGAGGTGCGAGGTTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAATTTCTCA 388
 Db 637 GATTTAGCCAGAGATTTCAGATTACAGAGCCGCTGTTTCTATGGCTTCAAATTTGCCA 696
 QY 389 TCGAGAAATTTCACTCAGAGATGTACAGTTTGTGTATAGACACTTACATCAGAGATCCCA 448
 Db 697 TGGAAAAACATACATTCTGAAAATGTATAGTCTTCTTATTGACACTTACATAAAGATCCCA 756
 QY 449 AGAAAAGGGAATTTTATTATTAATGCAATTTGAAACCATTCGCCCTATGTTAAGAAAAAGCAG 508
 Db 757 AAGAAAAGGGAATTTCTCTTCAATGCCATTTGAAACGATGCCCTTGTGTCAAGAAAGGAGCAG 816
 QY 509 ATTGGSCCTTGGATGGATAGCAGATAGAAAATCTACTTTTGGGAAAGAGTGTGTGCGCT 568
 Db 817 ACTGGCCTTGGCCTGGATTTGGGACAAAGAGGCTTACCTATGTGTGAACGTGTGTAGCCT 876

Db 1357 ACTT 1360

Search completed: October 30, 2005, 00:02:46
Job time : 442.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2005, 23:30:35 ; Search time 146 Seconds
(without alignments)
11801.366 Million cell updates/sec

Title: US-10-698-228-12

Perfect score: 1053

Sequence: 1 atggggcgaccggaaaggcc.....tcaccttgatgcagatttt 1053

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/prodata/1/ina/5A COMB.seq.*
- 2: /cgn2_6/prodata/1/ina/5B COMB.seq.*
- 3: /cgn2_6/prodata/1/ina/6A COMB.seq.*
- 4: /cgn2_6/prodata/1/ina/6B COMB.seq.*
- 5: /cgn2_6/prodata/1/ina/PCTUS COMB.seq.*
- 6: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1053	100.0	1053	4	US-10-019-733-12
2	1053	100.0	1081	4	US-10-019-733-4
3	1051.4	99.8	1053	4	US-10-019-733-2
4	1051.4	99.8	4955	4	US-10-019-733-3
5	580.6	55.1	2500	4	US-09-962-665-9
6	580.6	55.1	2500	4	US-09-963-333-9
7	580.6	55.1	2500	4	US-09-962-677-9
8	580	55.1	2500	4	US-09-023-655-1370
9	580	55.1	2500	4	US-09-949-016-145
10	578.4	54.9	2479	3	US-09-949-016-2025
11	439.4	41.7	481	3	US-08-905-223-125
12	429.8	40.8	14176	1	US-08-307-499-1
13	429.8	40.8	14176	1	US-08-307-499-14
14	429.8	40.8	14176	3	US-09-299-268-1
15	429.8	40.8	14176	3	US-09-299-268-14
16	384.8	36.5	1236	4	US-09-248-796A-3914
17	382.8	36.4	825	3	US-08-307-499-40
18	382.8	36.4	825	3	US-09-299-268-40
19	339.4	32.2	1112	4	US-08-136-743B-1
20	317.6	30.2	1083	4	US-09-248-796A-3915
21	216.6	30.1	419	3	US-08-905-223-129
22	260.8	24.8	656	3	US-09-328-111-791
23	195.2	18.5	449	4	US-09-270-767-14670
24	169.2	16.1	366	3	US-08-307-499-13
25	169.2	16.1	366	3	US-09-299-268-13
26	141.4	13.4	373	4	US-09-401-064-337
27	139.8	13.3	373	4	US-09-401-064-319

c	28	115.2	10.9	366	4	US-09-401-064-307	Sequence 307, App
	29	112.4	10.7	297	4	US-09-313-294A-4435	Sequence 4435, Ap
	30	93.8	8.9	11820	4	US-09-949-016-11887	Sequence 11887, A
	31	93.8	8.9	11826	4	US-09-949-016-13767	Sequence 13767, A
	32	92	8.7	601	4	US-09-949-016-20615	Sequence 20615, A
	33	92	8.7	601	4	US-09-949-016-69028	Sequence 69028, A
	34	76	7.2	276	4	US-09-313-294A-3437	Sequence 3437, Ap
	35	71.8	6.8	304	4	US-09-313-294A-7143	Sequence 7143, Ap
	36	68	6.5	276	4	US-09-313-294A-4638	Sequence 4638, Ap
	37	61	5.8	1053	4	US-09-902-540-2666	Sequence 2666, Ap
	38	61	5.8	13706	4	US-09-902-540-1124	Sequence 1124, Ap
	39	60.8	5.8	1141	4	US-09-806-708B-22	Sequence 22, Appl
	40	60	5.7	601	4	US-09-949-016-69031	Sequence 69031, A
	41	54.4	5.2	288	4	US-09-313-294A-6108	Sequence 6108, Ap
	42	53.8	5.1	276	4	US-09-313-294A-3244	Sequence 3244, Ap
	43	50	4.7	1230025	4	US-09-198-452A-1	Sequence 1, Appl
	44	50	4.7	1230230	4	US-09-438-185A-1	Sequence 1, Appl
	45	47	4.5	124884	4	US-09-661-596A-76	Sequence 76, Appl

ALIGNMENTS

RESULT 1

US-10-019-733-12

; Sequence 12, Application US/10019733

; Patent No. 6682917

; GENERAL INFORMATION:

; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA

; TITLE OF INVENTION: New Protein and its DNA

; FILE REFERENCE: 2619W00P

; CURRENT APPLICATION NUMBER: US/10/019,733

; CURRENT FILING DATE: 2001-12-28

; PRIOR APPLICATION NUMBER: JP 11-181131

; PRIOR FILING DATE: 1999-06-28

; PRIOR APPLICATION NUMBER: JP 11-192391

; PRIOR FILING DATE: 1999-07-06

; PRIOR APPLICATION NUMBER: JP 2000-017770

; PRIOR FILING DATE: 2000-01-21

; NUMBER OF SEQ ID NOS: 14

; SEQ ID NO 12

; LENGTH: 1053

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION:

US-10-019-733-12

Query Match 100.0%; Score 1053; DB 4; Length 1053;
Best Local Similarity 100.0%; Pred. No. 6.8e-309;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGGGGCACC	CGGAAAGG	CGCGCGGCTGGATCAGGATGAGATCATCTTCA	60
DB	1	ATGGGGCACC	CGGAAAGG	CGCGCGGCTGGATCAGGATGAGATCATCTTCA	60
QY	61	GACACCAAC	GAAAGTGA	ATAAGTCAATGAAGACCACTCTCTAAGAAAGAGTTCTCGC	120
DB	61	GACACCAAC	GAAAGTGA	ATAAGTCAATGAAGACCACTCTCTAAGAAAGAGTTCTCGC	120
QY	121	CGGTTGTG	TATCTTTT	CCAATCCAGTACCCCTGATTTTGGAAAAATGTATAA	180
DB	121	CGGTTGTG	TATCTTTT	CCAATCCAGTACCCCTGATTTTGGAAAAATGTATAA	180
QY	181	GCTTCTCT	CGGACACG	AGAGAGTTGACTTCAAGGATCTCCCTCACTGGAACAG	240
DB	181	GCTTCTCT	CGGACACG	AGAGAGTTGACTTCAAGGATCTCCCTCACTGGAACAG	240
QY	241	CTTAAAGC	AGATGAGA	AGTACTTCACTCTTACAGTTTTCGAGCCAGTGAT	300
DB	241	CTTAAAGC	AGATGAGA	AGTACTTCACTCTTACAGTTTTCGAGCCAGTGAT	300
QY	301	GGAATTGT	AAATGAAA	TTTGGTGGAGCGCTTTAGTCAGGAGGTGAGGTTCCAGAGCT	360

```
Db 301 GGAAATTGTAATAAGAAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
Qy 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAAGATGTTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAAGATGTTTCACTCAGAGATGTACAGTTTG 420
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAAATCAATTTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAAATCAATTTGAA 480
Qy 481 ACCATGCCCTATGTTTAAGAAAAAGCAGATTGGGCTTTGCGATGGATAGCAGATAGAAAA 540
Db 481 ACCATGCCCTATGTTTAAGAAAAAGCAGATTGGGCTTTGCGATGGATAGCAGATAGAAAA 540
Qy 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 600
Qy 601 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGTCTTTATGCCAGGACTCACTTTTCC 660
Db 601 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGTCTTTATGCCAGGACTCACTTTTCC 660
Qy 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGGCTGTGTTCCAA 720
Db 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGGCTGTGTTCCAA 720
Qy 721 TACTTAGTAATAAGCCCTTCAGAAAAGGGTTCAGGAGATCATTTGTTGATGCTGTCAA 780
Db 721 TACTTAGTAATAAGCCCTTCAGAAAAGGGTTCAGGAGATCATTTGTTGATGCTGTCAA 780
Qy 781 ATTGAGCAGGAGTTTAAACAGAACCTTTGCCAGTTGGCTTCATTGGGAATGAATTCGATT 840
Db 781 ATTGAGCAGGAGTTTAAACAGAACCTTTGCCAGTTGGCTTCATTGGGAATGAATTCGATT 840
Qy 841 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGGAATTCGATTCTCA 900
Db 841 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGGAATTCGATTCTCA 900
Qy 901 AAGTTTTTTCAGGCGAAGAAATCCTTTGATTTTATGAAAACATTTCTTTAGAGGAAAA 960
Db 901 AAGTTTTTTCAGGCGAAGAAATCCTTTGATTTTATGAAAACATTTCTTTAGAGGAAAA 960
Qy 961 ACAAAATTTCTTGAGAAACGAGTTTCAGAGTATCAGCGTTTGTGAGTTTATGCGAGAAC 1020
Db 961 ACAAAATTTCTTGAGAAACGAGTTTCAGAGTATCAGCGTTTGTGAGTTTATGCGAGAAC 1020
Qy 1021 ACAGATAACGCTTCCACCTTGGATGCAGATTTT 1053
Db 1021 ACAGATAACGCTTCCACCTTGGATGCAGATTTT 1053
```

RESULT 2

```
US-10-019-733-4
; Sequence 4, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 4
; LENGTH: 1081
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-4
Query Match 100.0%; Score 1053; DB 4; Length 1081;
Best Local Similarity 100.0%; Pred. No. 6.9e-309;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGGGACCCCGAAAGCGCGGCGGGCTGGATCAGGATCAGAGATCATCTTCA 60
Db 20 ATGGGCGAGCCCGAAAGCGCGGCGGGCTGGATCAGGATCAGAGATCATCTTCA 79
Qy 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTCTCGC 120
Db 80 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTCTCGC 139
Qy 121 CGGTTTGTCACTTTTCCAAATCCAGTACCTCTGATATTTTGAAAAATGTATAAACAGGCACAG 180
Db 140 CGGTTTGTCACTTTTCCAAATCCAGTACCTCTGATATTTTGAAAAATGTATAAACAGGCACAG 199
Qy 181 GCTTCTCTTCGACACAGCAGAGAGGTGTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 240
Db 200 GCTTCTCTTCGACACAGCAGAGAGGTGTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 259
Qy 241 CTTAAAGCAGATGAGAAGTACTTCTCTCATCTTACATCTTAGCCCTTTTTCGAGCCAGTGAT 300
Db 260 CTTAAAGCAGATGAGAAGTACTTCTCTCATCTCTCATCTTAGCCCTTTTTCGAGCCAGTGAT 319
Qy 301 GGAATTGTAATGAAAAATTTGGTGGAGCGCTTTTAGTCAGGAGTGCAGGTTCCAGAGGCT 360
Db 320 GGAATTGTAATGAAAAATTTGGTGGAGCGCTTTTAGTCAGGAGTGCAGGTTCCAGAGGCT 379
Qy 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCTGAGAAATGTTTCACTCAGAGATGTACAGTTTG 420
Db 380 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAAGATGTTTCACTCAGAGATGTACAGTTTG 439
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTTGAA 480
Db 440 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTTGAA 499
Qy 481 ACCATGCCCTATGTTTAAGAAAAAGCAGATTGGGCTTTCGATGGATAGCAGATAGAAAA 540
Db 500 ACCATGCCCTATGTTTAAGAAAAAGCAGATTGGGCTTTCGATGGATAGCAGATAGAAAA 559
Qy 541 TCTACTTTTGGGAAAGAGTGTGGCTTTTGTCTGTGTAGAGGAGTTTCTTCTCAGGA 600
Db 560 TCTACTTTTGGGAAAGAGTGTGGCTTTTGTCTGTGTAGAGGAGTTTCTTCTCAGGA 619
Qy 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGGTCTTTATGCCAGGACTCACTTTTTC 660
Db 620 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGGTCTTTATGCCAGGACTCACTTTTTC 679
Qy 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTTGCCTGATGTTCCAA 720
Db 680 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTTGCCTGATGTTCCAA 739
Qy 721 TACTTAGTAATAAGCCCTTCAGAAAAGGGTTCAGGAGATCATTTGTTGATGCTGTCAA 780
Db 740 TACTTAGTAATAAGCCCTTCAGAAAAGGGTTCAGGAGATCATTTGTTGATGCTGTCAA 799
Qy 781 ATTGAGCAGGAGTTTAAACAGAACCTTTGCCAGTTGGCTTCATTCGGAATGAATTCGATT 840
Db 800 ATTGAGCAGGAGTTTAAACAGAACCTTTGCCAGTTGGCTTCATTCGGAATGAATTCGATT 859
Qy 841 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGGAATTCGATTCTCA 900
Db 860 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGGAATTCGATTCTCA 919
Qy 901 AAGTTTTTTCAGGCGAAGAAATCCTTTTGTATTTTATGAAAAACATTTCTTTAGAGGAAAA 960
Db 920 AAGTTTTTTCAGGCGAAGAAATCCTTTTGTATTTTATGAAAAACATTTCTTTAGAGGAAAA 979
Qy 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCGAGTTTATGCGAGAAC 1020
```

Db 980 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTATGGCAGAAACC 1039
QY 1021 ACAGATAACGCTCTTACCTTGGATGAGATTTT 1053
Db 1040 ACAGATAACGCTCTTACCTTGGATGAGATTTT 1072

RESULT 3

US-10-019-733-2
; Sequence 2, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-2

Query Match 99.8%; Score 1051.4; DB 4; Length 1053;
Best Local Similarity 99.9%; Pred. No. 2.1e-308;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAGGCGGAGCGCGGCTGGATCAGGATGAGATCATCTTCA 60
Db 1 ATGGGCGACCCGGAAGGCGGAGCGCGGCTGGATCAGGATGAGATCATCTTCA 60
QY 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGGCCACTCTTAAGAAAGAGTTCTCGC 120
QY 121 CGGTTTGTCATCTTTCCAAATCCAGTACCTCGATTTTGGAAAAATGTATAACAGGCAAG 180
Db 121 CGGTTTGTCATCTTTCCAAATCCAGTACCTCGATTTTGGAAAAATGTATAACAGGCAAG 180
QY 181 GCTTCCTCTGGACAGCAGAGAGGTTGACTTATCAAGAGTCTCCCTCACTGGAAACAAG 240
Db 181 GCTTCCTCTGGACAGCAGAGAGGTTGACTTATCAAGAGTCTCCCTCACTGGAAACAAG 240
QY 241 CTTAAAGCAGATGAGAAGTACTCTCTCATCTTACATCTTAGCCCTTTTTCAGCCAGTGAT 300
Db 241 CTTAAAGCAGATGAGAAGTACTCTCTCATCTTAGCCCTTTTTCAGCCAGTGAT 300
QY 301 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCGAGTTCCAGAGGCT 360
Db 301 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCGAGTTCCAGAGGCT 360
QY 361 CGCTGTTCTATGGCTTTTCAAAATTCATCGAAGATGTTTCACTCAGAGATGTACAGTTG 420
Db 361 CGCTGTTCTATGGCTTTTCAAAATTCATCGAAGATGTTTCACTCAGAGATGTACAGTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTTGAA 480
QY 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGGCTTTCGATGGATAGCAGATAGAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGGCTTTCGATGGATAGCAGATAGAAA 540

QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGATTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGATTTTCTTCTCAGGA 600
QY 601 TCTTTTGGCTGCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCGCAGACTCATCTTTTCC 660
Db 601 TCTTTTGGCTGCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCGCAGACTCATCTTTTCC 660
QY 661 AATGAACACTCATCAGCAGAGATGAAGGACTTCACGTGACCTTTCCTTCCCTGATGTTCAA 720
Db 661 AATGAACACTCATCAGCAGAGATGAAGGACTTCACGTGACCTTTCCTTCCCTGATGTTCAA 720
QY 721 TACTTAGTAAATAAGCCTTCAGAGAAGAGGTTCAGGAGAGTCAATTGTTGATGCTGTCAAA 780
Db 721 TACTTAGTAAATAAGCCTTCAGAGAAGAGGTTCAGGAGAGTCAATTGTTGATGCTGTCAAA 780
QY 781 ATTGAGCAGAGTTTTAAACAGAAAGCCTTGCCAGTTGGCCTCATTTGGAATGAATTCATT 840
Db 781 ATTGAGCAGAGTTTTAAACAGAAAGCCTTGCCAGTTGGCCTCATTTGGAATGAATTCATT 840
QY 841 TTGATGAAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTTCTCA 900
Db 841 TTGATGAAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTTCTCA 900
QY 901 AAGGTTTTTTCAGGCAGAGAAATCCTTTTGTATTTATGAAAAACATTTCTTTAGAAAGAAA 960
Db 901 AAGGTTTTTTCAGGCAGAGAAATCCTTTTGTATTTATGAAAAACATTTCTTTAGAAAGAAA 960
QY 961 ACAAAATTTCTTTGAGAAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAAAACC 1020
Db 961 ACAAAATTTCTTTGAGAAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAAAACC 1020
QY 1021 ACAGATAACGCTCTTACCTTGGATGAGATTTT 1053
Db 1021 ACAGATAACGCTCTTACCTTGGATGAGATTTT 1053

RESULT 4

US-10-019-733-3
; Sequence 3, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 3
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-3

Query Match 99.8%; Score 1051.4; DB 4; Length 4955;
Best Local Similarity 99.9%; Pred. No. 5.1e-308;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAGGCGGAGCGCGGCTGGATCAGGATGAGATCATCTTCA 60
Db 245 ATGGGCGACCCGGAAGGCGGAGCGCGGCTGGATCAGGATGAGATCATCTTCA 304
QY 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTTCTCGC 120
Db 305 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTTCTCGC 364

121 CGGTTTGTTCATCTTTCCAAATCAGTACCTGATATTTGGAAAATGTATAAAGGACACAG 180
Db CGGTTTGTTCATCTTTCCAAATCAGTACCTGATATTTGGAAAATGTATAAAGGACACAG 424
QY 181 GCTTCTCTTCTGGACAGCAGAGAGGTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 240
Db 425 GCTTCTCTTCTGGACAGCAGAGAGGTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 484
QY 241 CTTAAGCAGATGAGAGTACTTCTCTCATCTTACATCTTAGCCCTTTTTCAGCCAGTGAT 300
Db 485 CTTAAGCAGATGAGAGTACTTCTCTCATCTTACATCTTAGCCCTTTTTCAGCCAGTGAT 544
QY 301 GGAATTGTAAATGAAAATTTGGTGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 360
Db 545 GGAATTGTAAATGAAAATTTGGTGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 604
QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAAGTGTTCACCTCAGAGATGTACAGTTTG 420
Db 605 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAAGTGTTCACCTCAGAGATGTACAGTTTG 664
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA 724
QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATTCGGCCCTTCGGATGGATAGCAGATGAAA 540
Db 725 ACCATGCCCTATGTTAAGAAAAGCAGATTCGGCCCTTCGGATGGATAGCAGATGAAA 784
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGTCTGTAGAGAGGTTTTCTTCTCAGGA 600
Db 785 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGTCTGTAGAGAGGTTTTCTTCTCAGGA 844
QY 601 TCTTTTGTCTATATTCGGCTTAAGAAAGAGGTCATTATGCCAGGACTCATTTTTC 660
Db 845 TCTTTTGTCTATATTCGGCTTAAGAAAGAGGTCATTATGCCAGGACTCATTTTTC 904
QY 661 AATGAACCTCATCAGCAGATCAAGACCTTCACTGTGACTTTTGTCTGTGCTGTGATTC 720
Db 905 AATGAACCTCATCAGCAGATCAAGACCTTCACTGTGACTTTTGTCTGTGCTGTGATTC 964
QY 721 TACTTAGTAAATAGCCCTTCAGAAAGAAAGGTCAGGAGATCATTTGTGATGCTGCAAA 780
Db 965 TACTTAGTAAATAGCCCTTCAGAAAGAAAGGTCAGGAGATCATTTGTGATGCTGCAAA 1024
QY 781 ATTGAGCAGGATTTTAAAGAAAGCTTGCAGTTGGCTCATTTGGAATGAATTGCATT 840
Db 1025 ATTGAGCAGGATTTTAAAGAAAGCTTGCAGTTGGCTCATTTGGAATGAATTGCATT 1084
QY 841 TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTTACTTGTGGAATTTGATTTCTCA 900
Db 1085 TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTTACTTGTGGAATTTGATTTCTCA 1144
QY 901 AAGGTTTTTTCAGCAGAAAATCCTTTTGTATTTTATGAAAAACATTTCTTTAGAAAGAAA 960
Db 1145 AAGGTTTTTTCAGCAGAAAATCCTTTTGTATTTTATGAAAAACATTTCTTTAGAAAGAAA 1204
QY 961 ACAATTTCTTTCAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020
Db 1205 ACAATTTCTTTCAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1264
QY 1021 ACAGATACGTTCTTACCTTGGATGAGATTTT 1053
Db 1265 ACAGATACGTTCTTACCTTGGATGAGATTTT 1297

RESULT 5
US-09-962-665-9
; Sequence 9, Application US/09962665
; Patent No. 6537759
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: POLY(POLYGLUTAMATE SYNTHETASE GENE SEQUENCE

; TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING THE
; TITLE OF INVENTION: TREATMENT OF DISEASE
; FILE REFERENCE: 11926-015004
; CURRENT APPLICATION NUMBER: US/09/962,665
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 128, 1464
; OTHER INFORMATION: n = g or a
; NAME/KEY: misc_feature
; LOCATION: 189
; OTHER INFORMATION: n = t or g
; NAME/KEY: misc_feature
; LOCATION: 524
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc_feature
; LOCATION: 1399
; OTHER INFORMATION: n = t or a
; NAME/KEY: misc_feature
; LOCATION: 1636, 1738, 2259
; OTHER INFORMATION: n = c or t
; US-09-962-665-9

Query Match 55.1%; Score 580.6; DB 4; Length 2500;
Best Local Similarity 75.1%; Pred. No. 2,2e-165;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;
QY 89 ATCAAGAGCCACTCTTAAGAAAGAGTCTCGCGGTTTGTCACTTTCCCAATCCAGTACC 148
Db 397 AGGATAGCGCGTGTGAGAGAAAACCCCGCGCTTTGTCACTTTCCCAATCCAGTACC 456
QY 149 CTGATATTTGAAAAATGTATAAACAGGCACAGGCTTCTTCTGGACAGCAGAGGTTG 208
Db 457 ATGATATCTGGCAGATGTATAAGAAAGCAGAGGCTTCTTTTGACCCCGAGGAGTTG 516
QY 209 ACTTATCAAGAGATCTCCCTCACTGGAAACAAGCTTAAAGCAGATGAGAACTTCACTCT 268
Db 517 ACCTCTCNAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTATAT 576
QY 269 CTCACATCTTAGCCTTTTTCAGCCAGTATGGAATGTAAATGAAAATTTGGTGAGC 328
Db 577 CCATGTTCTGGCTTTCTTTTGACAGCAAGCATGGCATAGTAAATGAAAACCTTTGGTGGAGC 636
QY 329 GCTTTAGTCAGGAGGTGCGAGGTTCCAGAGGCTCGCTGTTTCTATGGCTTTTCAAAATTTCTCA 388
Db 637 GATTAGCCNAGAGTTTCAGATTACAGAGCCCGCTGTTTCTATGGCTTTCCAAATTTGCCA 696
QY 389 TCGAGATGTTCTACTCAGAGATGTACAGTTTGTGTATAGACACTTATCAGAGATCCCA 448
Db 697 TGGAAAACATACATTCGAAATGTATAGTCTTCTTATTGACACTTACATAAAGATCCCA 756
QY 449 AGAAAAGGAATTTTATTTAATGCAATTGAAAACCATGCCCTATGTTAAGAAAAAGCAG 508
Db 757 AAGAAAAGGAATTTTCTCTTCAATGCCATTTGAAACGATGCCCTTGTGTCAAGAAAGAGGCG 816
QY 509 ATTGGGCTTGGATGGATGAGATAGAAAATCTACTTTTGGGAAAGAGTGTGGCT 568


```
Db 817 ACTGGGCTTGGCGCTGGATTGGGGAACAAAGAGCTACCTATGCTGAACGTTGTAGCCT 876
Qy 569 TTGCTGCTCTAGAAGAGTTTCTTCTCAGGATCTTTTCTGCTGCTATATTCTGCTAAAGA 628
Db 877 TTGCTGCTGAGGAGCAATTTCTTTCCGGTCTTTTGGCTGCTGATATTCTGCTCAAGA 936
Qy 629 AGAGAGGCTTTATGCGGAGCTCACCTTTTCCAAATCACTATCAGCAGAGATGAAGGAC 688
Db 937 AACGAGGACTGATGCTGCGCTCACATTTTCTAAAGAACTTATAGCAGAGATGAGGTT 996
Qy 689 TTCACTGTGACTTTGCTTGGCTGATGTTCCAAATCACTATGTAATAAGCCTTCAGAAAGAA 748
Db 997 TACACTGTGATTTTGTCTGCTGATGTTTCAAAACACCTGTCACAAACCATCGAGGAGA 1056
Qy 749 GGGTCAGGAGATCATTTGTTGATGCTGTCAAAATTGAGCAGAGTTTAAACAGAGCCT 808
Db 1057 GAGTAAGAGAAATAATATCAATGCTTTCGGATAGAACAGAGTTCTCTCACTGAGCCT 1116
Qy 809 TGCCAGTTGGCCTCATTTGGAATGAATTGCAATTTTGATGAAACAGTACATTTGATTTGTAG 868
Db 1117 TGCTGTGAGCTCATTTGGATGAATTGCACTCTATGAAAGCAATACATTTGATTTGTGG 1176
Qy 869 CTGACAGATTACTTTGTGGAATTTGGAATTTCTCAAAGTTTTCAGGAGAGAAATCTCTTTG 928
Db 1177 CAGACAGACTTATGCTGGAATCGGTTTTCAGCAAGTTTTCAGAGTAGAGAACCCATTG 1236
Qy 929 ATTTTATGGAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACAGATTTTCAG 988
Db 1237 ACTTTATGGAATATTTTCACTGGAAGGAAAGACTAACTTCTTTGAGAGAGATGAGCG 1296
Qy 989 AGTATCAGGCTTTTGCAGTTATGCGAGATGCGAGAAACACAGATAACTTCTCACTTTGGATG 1048
Db 1297 AGTATCAGAGATGAGGATGATGTCAGTCCAAACAGAGAAATCTTTTACCTTTGGATGCTG 1356
Qy 1049 ATTT 1052
Db 1357 ACTT 1360
```

RESULT 6

```
US-09-963-333-9
; Sequence 9, Application US/09963333
; Patent No. 6664062
; GENERAL INFORMATION:
; APPLICANT: Stancon, Jr., Vincent P.
; TITLE OF INVENTION: THYMIDINE SYNTHASE GENE SEQUENCE VARIANCES
; TITLE OF INVENTION: HAVING UTILITY IN DETERMINING THE TREATMENT
; TITLE OF INVENTION: OF DISEASE
; FILE REFERENCE: 11926-015002
; CURRENT APPLICATION NUMBER: US/09/963,333
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 128, 1464
; OTHER INFORMATION: n = g or a
; NAME/KEY: misc_feature
; LOCATION: 189
```

```
; OTHER INFORMATION: n = t or g
; NAME/KEY: misc_feature
; LOCATION: 524
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc_feature
; LOCATION: 1399
; OTHER INFORMATION: n = t or a
; NAME/KEY: misc_feature
; LOCATION: 1636, 1738, 2259
; OTHER INFORMATION: n = c or t
; US-09-963-333-9
```

```
Query Match 55.1%; Score 580.6; DB 4; Length 2500;
Best Local Similarity 75.1%; Pred. No. 2.2e-165;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

Qy 89 ATGAAGAGCCTCTCTAAGAAAGAGTTCTCGCGGTTTGTTCATCTTTCCAAATCCAGTACC 148
Db 397 AGGATGAGCGCTGCTGAGAGAAAACCCCGCGCTTTGTCACTTCTCCCATCGAGTACC 456
Qy 149 CTGATATTTGGAAAATGTATAAACAGCACAGGCTTCTCTGACAGCAGAGAGGTTG 208
Db 457 ATGATATCTGGCAGATGTATAAGAGGACAGAGGCTTCTTTTGGACCGCGAGAGGTTG 516
Qy 209 ACTTATCAAGGATCTCCCTCACTGGAACAGCTTAAAGCAGATGAGAAGTACTTCTCT 268
Db 517 ACCTCTCAAGGACATTCAGACTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT 576
Qy 269 CTCACATCTTAGCCTTTTTCAGCCAGTGTGGAATTTGAAATGAAATTTGGTGGAGC 328
Db 577 CCCATGTTCTGCTTTTTCAGCAAGCGATGCGATAGTAATGAAACTTGGTGGAGC 636
Qy 329 GTTTAGTCAGGAGTGCGAGTTCCAGAGCTCGCTGTTTCTATGCTTTCAAAATCTCA 388
Db 637 GATTTAGCCAAAGATTTCAGATTACAGAGCCGCTGTTTCTATGCTTCCAAATGCCA 696
Qy 389 TCGAGNATGTTCACTCAGAGATGTACAGTTTGTGTATAGACACTTACATCAGAGATCCA 448
Db 697 TGGAAAAACATACATTTCTGAAATGTATAGTCTTCTTATGACACTTACATAAAGATCCA 756
Qy 449 AGAAAGAGGAATTTTATTTAATGCAATTTGAAACCATGCCCTATGTTAAAGAAAAACGAG 508
Db 757 AGAAAGAGGAATTTCTTCAATGCCATTTGAAACGATGCCCTTGTGTCAGAGAGAGCAG 816
Qy 509 ATTGGCCCTTGGCAGATAGCAGATAGAAAATCTACTTTTGGGGAAGAGTGGTGGCCT 568
Db 817 ACTGGCCCTTGGCTGATTTGGGACAAAGAGGCTACCTATGTTGAACGTTGTGTAGCCT 876
Qy 569 TTGCTGCTGTAGAGAGTTTCTTCTCAGGATCTTTTGTGCTGTATATTTCTGCTAAAGA 628
Db 877 TTGCTGCTGTAGAGAGCAATTTCTTTTCCGGTTTCTTTTGGCTGATATTTCTGGCTCAAGA 936
Qy 629 AGAGAGGCTTATGCGGAGCTCACCTTTTCCAAATGAACCTCATCAGCAGAGATGAAGGAC 688
Db 937 AACGAGGACTGATGCTGCGCTCACATTTTCTAATGAACCTTATAGCAGAGATGAGGTT 996
Qy 689 TTCCTGTGACTTTGCTTGGCTGATGTTCCAAATCACTTAGTAAATAAGCCTTCAGAGAAA 748
Db 997 TACACTGTGATTTGCTTGGCTGATGTTCAAAACACCTGGTACACAAACCATCGAGGAGA 1056
Qy 749 GGGTCAGGAGATCATTTGTTGATGCTGTCAAAATTGAGCAGAGGATTTTAAACAGAGCCT 808
Db 1057 GAGTAAGAGAAATAATATCAATGCTTTCGGATAGAAACAGGAGTTCTCTCACTGAGCCT 1116
Qy 809 TGCCAGTTGGCCTCATTTGGAATGAATTGCAATTTTGATGAAACAGTACATTTGATTTGTAG 868
Db 1117 TGCTGTGAGCTCATTTGGATGAATTGCACTCTAATGAAAGCAATACATTTGATTTGTGG 1176
Qy 869 CTGACAGATTACTTTGTGGAATTTGGAATTTCTCAAAGTTTTCAGGAGAGAAATCTCTTTG 928
Db 1177 CAGACAGACTTATGCTGGAATCGGTTTTCAGCAAGTTTTCAGAGTAGAGAACCCATTG 1236
Qy 929 ATTTTATGGAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACAGATTTTCAG 988
```

Db 1237 ACTTATGAGATATTTTCACTGGAGGAAAGACTAACTCTTTTGGAGAGAGTAGGCG 1296
QY 989 AGTATCAGCGTTTTCAGTTATGGCAGAAACACAGATAACGCTTTCACCTTGGATGCGAG 1048
Db 1297 AGTATCAGAGGATGGAGTGATGTCAAGTCCAACAGAGAATTTCTTTTACCTTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 7

US-09-962-677-9
; Sequence 9, Application US/09962677
; Patent No. 6759200
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: THYMIDINE PHOSPHORYLASE GENE SEQUENCE
; TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING
; TITLE OF INVENTION: THE TREATMENT OF DISEASE
; FILE REFERENCE: 11926-015003
; CURRENT APPLICATION NUMBER: US/09/962,677
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 128_1464
; OTHER INFORMATION: n = g or a
; NAME/KEY: misc_feature
; LOCATION: 189
; OTHER INFORMATION: n = t or g
; NAME/KEY: misc_feature
; LOCATION: 524
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc_feature
; LOCATION: 1399
; OTHER INFORMATION: n = t or a
; NAME/KEY: misc feature
; LOCATION: 1636_1738, 2259
; OTHER INFORMATION: n = c or t
US-09-962-677-9

Query Match 55.1%; Score 580.6; DB 4; Length 2500;
Best Local Similarity 75.1%; Pred. No. 2.2e-165;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;
QY 89 ATGAGAGCCATCTCTAAGAAAGAGTTCTGCGCGGTTTGTCTATCTTTTCCAAATCCAGTACC 148
Db 397 AGGATGAGCGCTGTGTGAGAGAAACCCCGCGCTTTGTCTATCTTCCCATCGAGTACC 456
QY 149 CTCGATATTTGGAATATGATAACAGGCACAGGCTTCTCTGGACAGCAGAGAGGTTG 208
Db 457 ATGATATCTGGCAGATGTATAGAGGCGAGGCTTCTCTTTTGGACCGCGGAGGAGTTG 516
QY 209 ACTTATCAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAGTACTTCTCATCT 268
Db 517 ACCTCTCNAGGACATTGACACTGGGAATCTCTGAAACCCGAGGAGAGATATTTATAT 576

QY 269 CTCACATCTTAGCCTTTTTCAGCCAGTGATGAATCTTAATGAAATTTTGGTGGAGC 328
Db 577 CCATGTTCTTGGCTTTTTCAGCAAGCGATGGCATGTAATGAAATTTTGGTGGAGC 636
QY 329 GCTTTAGTCAGGAGTGCAGAGGCTTCAGAGGCTCGCTGTTTCTATGGCTTTCAAATTTCTCA 388
Db 637 GATTTAGCCAGAGTTTCAGATTACAGAACCCGCTGTTTCTATGGCTTTCCAATTTGCCA 696
QY 389 TCGAGAAATGTTCACTCAGAGATGTACAGTTTGTGTGATAGACACTTATCATCAGAGATCCCA 448
Db 697 TGGAAAAACATACATTCTGAAATGTATAGTCTTCTTATTCACACTTACATAAAAAAGATCCCA 756
QY 449 AGAAAGGGAATTTTATTTAATGCAATTTGAAACCATGCCCTATGTTAAGAAAAAAGCAG 508
Db 757 AAGAAAGGGAATTTCTCTTCAATGCAATTTGAAACGATGCTTGTGTCAAGAAAGAGCAG 816
QY 509 ATTGGGCTTGGCATGGATAGCAGATAGAAAAATCTACTTTTGGGGAAGAGTGGTGGCCT 568
Db 817 ACTGGGCTTGGCTGGATTGGGACAAAGAGCTACCTATGTTGAACGTTGTGTAGCCT 876
QY 569 TTGCTGCTGTAGAAGAGTTTCTTCTCAGGATCTTTTGTGCTGTATATTTCTGCTAAAGA 628
Db 877 TTGCTGCTGTAGAAGGCAATTTCTTTCGGTCTTCTTTCGGTCTGATATTTCTGCTCAAGA 936
QY 629 AGAGAGGCTTATGCCAGGACTCACTTTTCCAAATCACTATCATCAGCAGAGATGAAGGAC 688
Db 937 AAGCAGGACTGATGCTGCTGCTTCACTTTTCTAATGAACCTTATTAGCAGAGATGAGGGTT 996
QY 689 TTTCACTGTGACTTTGCTTGCCTGATGTTCCAAATACCTTAGTAAATAAGCCCTTCAGAAAGAA 748
Db 997 TACACTGTGATTTTGTCTGCTGATGTTCCAAACCTGTTACACAAACCATCGAGGAGA 1056
QY 749 GGGTCAGGAGATCATTTGTTGATGCTTCAAAATTCAGCAGGAGTTTAAACAGAGCCT 808
Db 1057 GAGTAAGAGAAATAATTATCAATGCTTTCGGATAGAACAGGAGTTTCTTCACTGAGGCT 1116
QY 809 TGCCAGTTGGCCTCATTTGAATGAATGCAATTTGATGAACAGTACATGAGTTGTAG 868
Db 1117 TGCCGTGTAAGCTCATTTGGGATGAATTCACCTTAATGAAGCAATACATTGAGTTGTGG 1176
QY 869 CTGACAGATTACTTTGGAACTTGGATTCTCAAGGTTTTCAGGAGAGAAATCTTTTG 928
Db 1177 CAGACAGACTTATGCTGGAATCTGGGTTTTCAGAGGTTTTCAGAGTAGAGAACCCATTG 1236
QY 929 ATTTTATGGAACATTTCTTTAGAGGAAAAACAAATTTCTTTTGAGAAACGAGTTTCAG 988
Db 1237 ACTTTATGGAATAATTTCACTGGAAGGAAAGACTAACTTCTTTTGAGAGAGTAGGCG 1296
QY 989 AGTATCAGCGTTTTCAGATTATGGCAGAAACCCAGATTAACGTTCTTCACTTGGATGCGAG 1048
Db 1297 AGTATCAGAGGATGGGAGTGATGTCAAGTCCAACAGAGAAATTTCTTTTACCTTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 8

US-09-023-655-1370
; Sequence 1370, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO

STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/023,655
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0001 US
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1370:
SEQUENCE CHARACTERISTICS:
LENGTH: 2500 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g36154
US-09-023-655-1370

Query Match 55.1%; Score 580; DB 4; Length 2500;
Best Local Similarity 75.1%; Pred. No. 3.4e-165;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

89 ATGAAGAGCCCTCTTAAGAAAGAGGTTCTGCGCGGTTTGTCATCTTTTCCAAATCCAGTACC 148
397 AGGATGAGCGCGTGTGAGAGAAACCCCGCGGTTTGTCATCTTTCCCATCGAGTACC 456
149 CTGATATTTGGAATATGATAAAGGAGGTTCTTCTGAGCAAGGTTTCTTCTGAGCAAGGTTG 208
457 ATGATATCTGGCAGATGATATAAGAGGAGGTTCTTCTTGGACCGCGGAGGTTG 516
209 ACTTATCAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAACTTCTCATCT 268
517 ACCCTCCAGGACATTCAGCATCTGGGATCCCTGAAACCCGAGGAGATATTTATAT 576
269 CTCACATCTTAGCCTTTTTCAGCAGTGATGGAATTTGTAATGAAATTTGGTGAGC 328
577 CCCATGTTCTGGCTTTCTTTCAGCAAGCGATGGCATAGTAAATGAAACTTTGGTGAGC 636
329 GCTTTAGTCAGAGTGCGAGGTTCCAGAGGCTCGCTGTTCTATGGCTTTCAATCTCA 388
637 GATTTAGCAAGAAGTTCAGATTAAGAGGCGCGCTGTTCTATGGCTTTCCAAATTTGCCA 696
389 TCGAAGATGTTTCACCTCAGAGATGTACAGTTTCTGATAGACACTTACATCAGAGATCCCA 448
697 TGGAAACATACATCTGAAATGATAGTCTTCTTATTTGACACTTACATAAAGATCCCA 756
449 AGAAAGGGAATTTTATTAATGCAATTTGAAACCATGCCCTPATGTTAAGAAAGGAGCAG 508
757 AAGAAAGGGAATTTCTCTTCAATGCCATTTGAAACGATGCTTGTGTCAGAAAGAGGCGAG 816
509 ATTGGCCTTGGCATGATAGATAGATAAATCTACTTTTGGGGAAGAGGTTGGCT 568
817 ACTGGCCTTGGCTTGGATTTGGGGAAGAGGCTACCTATGTTGAAACGTTGTAGCT 876
569 TTGCTGCTGTAGAAGGAGTTTCTTCTCAGGATCTTTTCTGCTATATTTCTGCTGCTAAAGA 628
877 TTGTCAGTGAAGGCAATTTCTTTTCGGTTCTTTTTCGGTCTGATATTTCTGGCTCAAGA 936

629 AGAGAGGCTCTTATGSCCAGGACTCACTTTTCCAAATGAACATCATCAGCAGAGATGAAGGAC 688
937 AACGAGACTGATGCGCTGGCCTCACATTTTCTAATGAACCTTATTAGCAGAGATGAGGGTT 996
689 TTCACCTGTGACTTTTGTCTGCTGATGTTTCAATATCTTAGTAAATAAGCCTTTCAGAAAGAA 748
997 TACACTGTGATTTTGTCTGCTGATGTTTCAACACCTGCTACACAAACCATCGGAGGAGA 1056
749 GGTTCAGGAGATCATTTGTTGATGCTCTCAAAATTTAGCAGGAGGTTTAAACAGAGCCT 808
1057 GAGTAAGAGAAATAAATTTATCAATGCTGTTTCGATAGAAACAGGAGTTCTCTCACTGAGGCT 1116
809 TGCAGTTGGCCTCATTTGGAATGAATTTGCAATTTTGTGAAACAGTACATTCAGTTTGTAG 868
1117 TGCCTGTGAAGCTCAITGGGATGAATTTGCACTCTAATGAAGCAATACATTTAGTTTGTGG 1176
869 CTGACAGATTACTTTGTGGAATCTTGGATTTCTCAAAAGGTTTTTTCAGGCGAGAAATCTCTTTG 928
1177 CAGCAGACTTATGCTGGAATCTGGTTTTCAGCAAGTTTTCAGAGTAGAGAACCATTTG 1236
929 ATTTATGGAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACAGTTTTCAG 988
1237 ACTTTATGGAGAATAATTTCACTGGAAGAAAGACTAACTTTCTTTGAGAAAGAGTAGGCG 1296
989 AGTATCAGCGTTTTCAGTTTATGCGCAAAACACAGATAAGCTTTTACCTTTGGATGCGAG 1048
1297 AGTATCAGAGGATGGGAGTGTATCAAGTCCACAGAGAAATTTCTTTTACCTTTGGATGCTG 1356
1049 ATTT 1052
1357 ACTT 1360

RESULT 9

US-09-949-016-145
; Sequence 145, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 145
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-145

Query Match 55.1%; Score 580; DB 4; Length 2500;
Best Local Similarity 75.1%; Pred. No. 3.4e-165;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;
89 ATGAAGAGCCCTCTTAAGAAAGAGGTTCTGCGCGGTTTGTCATCTTTTCCAAATCCAGTACC 148
397 AGGATGAGCGCGTGTGAGAGAAACCCCGCGGTTTGTCATCTTTCCCATCGAGTACC 456
149 CTGATATTTGGAATATGATAAAGGAGGTTCTTCTGAGCAAGGTTTCTTCTGAGCAAGGTTG 208
457 ATGATATCTGGCAGATGATATAAGAGGAGGTTCTTCTTGGACCGCGGAGGTTG 516
209 ACTTATCAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAACTTCTCATCT 268

Db 517 ACCTCTCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT 576
Qy 269 CTCACATCTTAGCCTTTTTCAGCCAGTGGATGTAATGAAATTTTGGTGAGC 328
Db 577 CCCATGTTCTGGCTTTTTCAGCAAGCGATGGCATAGTAATGAAATTTTGGTGAGC 636
Qy 329 GCTTTAGTCAGGAGTGCAGGTTTCAGAGGCTCGCTGTTTCTATGGCTTTCAAATTTCTCA 388
Db 637 GATTTAGCAAGAAGTTCAGATTCAGAGCCGCTGTTTCTATGGCTTCGAAATTTGCA 596
Qy 389 TCGAATATGTTCACTCAGAGATGTACAGTTTCTGATGACACTTACATCAGAGATCCCA 448
Db 697 TCGAATAACATACATCTGAAATGTATAGTCTTCTTATGACACTTACATAAAGATCCCA 756
Qy 449 AGAAAGGGAATTTTATTAATGCAATGGAACCATGCGCTATGTTAAGAAAGGAGCAG 508
Db 757 AAGAAAGGGAATTTCTCTCAATGCAATGGAACCATGCGCTTGTGTCAGAAAGAGGAGC 816
Qy 509 ATTGGGCTTGGATGGATAGAGATAGAAATCTACTTTTGGGAAAGAGTGGTGGCT 568
Db 817 ACTGGGCTTGGCTGGATGGGAAAGAGGCTACCTATGTTGAACTGTTGAGCT 876
Qy 569 TTGCTGCTAGAGGAGTTTCTTCTCAGGATCTTTTGGCTCTATATTTCTGGCTAAAGA 628
Db 877 TTGCTGCAAGTGAAGGCAATTTCTTTTTCGGTTCGATATTTCTGGCTCAAGA 936
Qy 629 AGAGAGTCTTATGCGAGGACTCCTTTTCAATGAACCTATCAGCAGAGATGAAGGAC 688
Db 937 AACGAGGACTGATGCTGCGCTCACAATTTCTTAATGAACCTTATAGCAGAGATGAGGTT 996
Qy 689 TTCACGTGACCTTCTGCTGATGTTCCCAATCTAGTAATTAAGCCTTCAGAAAGAA 748
Db 997 TACACTGTGATTTGCTGCTGATGTTCCCAATCTAGTAATTAAGCCTTCAGAAAGAA 1056
Qy 749 GGGTCAGGGAGATCATTTGATGCTGCAAAATTCAGCAGGAGTGTTCAGCAAGACCT 808
Db 1057 GAGTAAGAGAAATTAATCAATGCTGTCGGAATGAACAGGAGTTCCTCACTGAGGCT 1116
Qy 809 TGCCAGTTGGCTCATTTGGAATGAAATTTGATGAAACAGTACATTTGATGTTGAG 868
Db 1117 TGCTGTGAAGCTCATTTGGAGTAAATTCACCTAATGAAGCAATACATTTGAGTTGAG 1176
Qy 869 AGAGAGTCTTATGCGAGGACTCCTTTTTCAAATGAACTCATCAGCAGAGATGAAGGAC 688
Db 937 AACGAGGACTGATGCTGCGCTCACAATTTCTTAATGAACCTTATAGCAGAGATGAGGTT 996
Qy 689 TTTCACTGTGACTTTGCTGCTGATGTTCCCAATCTAGTAATTAAGCCTTCAGAAAGAA 748
Db 997 TACACTGTGATTTGCTGCTGATGTTCCCAATCTAGTAATTAAGCCTTCAGCAAGACCT 1056
Qy 749 GGGTCAGGGAGATCATTTGATGCTGCAAAATTCAGCAGGAGTGTTCAGCAAGACCT 808
Db 1057 GAGTAAGAGAAATTAATCAATGCTGTCGGAATGAACAGGAGTTCCTCACTGAGGCT 1116
Qy 809 TGCCAGTTGGCTCATTTGGAATGAAATTTGATGAAACAGTACATTTGATGTTGAG 868
Db 1117 TGCTGTGAAGCTCATTTGGAGTAAATTCACCTAATGAAGCAATACATTTGAGTTGAG 1176
Qy 869 CTGACAGATTTCTGTTGGAATCTTCAAGGTTTTCAGGCGAAGAAATTCCTTTT 928
Db 1177 CAGACAGACTTATGCTGGAACTGGGTTTTCAGGAGGTTTTCAGAGTGAAGACCCATTT 1236
Qy 929 ATTTTATGGAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACGAGTTTCAG 988
Db 1237 ACTTTATGGAATATTTCACTGGAAGGAAAGACTAACTTTCTTTGAGAAAGAGTAGGCG 1296
Qy 989 AGTATCAGCGTTTTCAGTTATGGCAGAAACACAGATTAACGCTTTACCTTTGATGAG 1048
Db 1297 AGTATCAGAGGATGGAGTGTCAAGTCCAAACAGAAATTTCTTTTACCTTTGGATGCTG 1356
Qy 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 10

US-09-949-016-2025
; Sequence 2025, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ FOR Windows Version 4.0
; SEQ ID NO 2025
; LENGTH: 2479
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-2025

Query Match 54.9%; Score 578.4; DB 4; Length 2479;
Best Local Similarity 75.0%; Pred. No. 1e-164;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;
Qy 89 ATGAAGAGCCATCTCTAAGAAAGAGTTCTCGCCGGTTTGTTCATCTTTCCCAATCCAGTACC 148
Db 397 AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTTGTTCATCTTTCCCATCGAGTACC 456
Qy 149 CTGATATTTGGAATATGTAACAGGCACAGCTTCTCTCGACAGCAGAGAGAGTTG 208
Db 457 ATGATATCTGGCAGATGTATAAGAGGCAGAGGCTTCTTTTGGACCGCAGAGAGTGG 516
Qy 209 ACTTATCAAGAGATCTCCCTCACTGGAACAAAGCTTAAAGCAGATGAGAACTACTTCATCT 268
Db 517 ACTCTCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTATAT 576
Qy 269 CTCACATCTTAGCCTTTTTCAGCCAGTGTAGGAATTTGTAATGAAATTTGGTGGAGC 328
Db 577 CCCATGTTCTGGCTTTTTCAGCAAGGATGGCATAGTAATGAAATTTGGTGGAGC 636
Qy 329 GCTTTAGTCAGAGGTCAGGTTCCAGAGGCTCGCTGTTCTATGGCTTTCAAATTTCTCA 388
Db 637 GATTTAGCCAAAGATTCAGATTTACAGAAAGCCGCTGTTCTATGGCTTCCAAATTTGCCA 696
Qy 389 TCGAATGTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAGATCCCA 448
Db 697 TCGAATAACATACATCTGAAATGTATAGTCTTCTTATGACACTTACATAAAGATCCCA 756
Qy 449 AGAAAGGGAATTTTATTTAATGCAATTTGAAACCATGCGCTATGTTAAGAAAGGAGCAG 508
Db 757 AAGAAAGGGAATTTCTCTTCAATGCCATTTGAAACGATGCTTGTGTCAAGAAAGGAGCAG 816
Qy 509 ATTGGGCTTGGATGGATAGAGATAGAAATCTACTTTTGGGAAAGAGTGGTGGCT 568
Db 817 ACTGGGCTTGGCTGGATGGGGAACAAAGAGGCTACCTATGTTGAAACGTTGTAGCCT 876
Qy 569 TTTGCTGCTAGAGGAGTTTCTCTCAGGATCTTTTGTCTGCTATATTTCTGGCTAAAGA 628
Db 877 TTGCTGCAAGTGAAGGCAATTTCTTTTCCGGTCTTTTGGCTGATATTTCTGGCTCAAGA 936
Qy 629 AGAGAGTCTTATGCGAGGACTCCTTTTTCAAATGAACTCATCAGCAGAGATGAAGGAC 688
Db 937 AACGAGGACTGATGCTGCGCTCACAATTTCTTAATGAACCTTATAGCAGAGATGAGGTT 996
Qy 689 TTTCACTGTGACTTTGCTGCTGATGTTCCCAATCTAGTAATTAAGCCTTCAGAAAGAA 748
Db 997 TACACTGTGATTTGCTGCTGATGTTCCCAATCTAGTAATTAAGCCTTCAGCAAGACCT 1056
Qy 749 GGGTCAGGGAGATCATTTGATGCTGCAAAATTCAGCAGGAGTGTTCAGCAAGACCT 808
Db 1057 GAGTAAGAGAAATTAATCAATGCTGTCGGAATGAACAGGAGTTCCTCACTGAGGCT 1116
Qy 809 TGCCAGTTGGCTCATTTGGAATGAAATTTGATGAAACAGTACATTTGATGTTGAG 868
Db 1117 TGCTGTGAAGCTCATTTGGAGTAAATTCACCTAATGAAGCAATACATTTGAGTTGAG 1176
Qy 869 CTGACAGATTTCTGTTGGAATCTTCAAGGTTTTCAGGCGAAGAAATTCCTTTT 928
Db 1177 CAGACAGACTTATGCTGGAACTGGGTTTTCAGGAGGTTTTCAGAGTGAAGACCCATTT 1236
Qy 929 ATTTTATGGAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACGAGTTTCAG 988

Db 1237 ACTTTAAGGAGATATTTTCACTGGAAGAAAGACTAACTTTCTTTGGAAGAGAGTAGGCG 1296
QY 989 AGTATCAGCGTTTTCAGTATTCGACAGAAACACAGATAACGCTTCACCTTCGATGCGAG 1048
Db 1297 AGTATCAGAGAGTGGAGTATGTCAGTCCAAACAGAGAAATCTTTTACCTTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 11

US-08-905-223-125
; Sequence 125, Application US/08905223
; Patent No. 6222029
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste D.
; APPLICANT: Duclert, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: 5' ESTs FOR SECRETED PROTEINS
; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobb, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Win95
SOFTWARE: Word

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/905,223
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Israel, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 125:
SEQUENCE CHARACTERISTICS:
LENGTH: 481 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: DOUBLE
TOPOLOGY: LINEAR
MOLECULE TYPE: CDNA
ORGANISM: Homo Sapiens
TISSUE TYPE: Brain
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 41..343
IDENTIFICATION METHOD: Von Heijne matrix
OTHER INFORMATION: score 4.4
OTHER INFORMATION: seq ISHILAFFAASDG/IV

US-08-905-223-125

Query Match 41.7%; Score 439.4; DB 3; Length 481;
Best Local Similarity 99.8%; Pred. No. 5.9e-123;
Matches 440; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGGACCCGGAAGCGGAGCGGGCTGGATCAGATCAGATCATTTCA 60
Db 41 ATGGGGACCCGGAAGCGGAGCGGGCTGGATCAGATCAGATCATTTCA 100
QY 61 GACACCAAGGAAGTGAATAAAGTCAATGAAGGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 101 GACACCAAGGAAGTGAATAAAGTCAATGAAGGCCACTCTTAAGAAAGAGTTCTCGC 160

QY 121 CGGTTTGTCTATCTTTCCAATCCAGTACCCTGATATTTGGAATAATGATATAACAGGCACAG 180
Db 161 CGGTTTGTCTATCTTTCCAATCCAGTACCCTGATATTTGGAATAATGATATAACAGGCACAG 220
QY 181 GCTTCCTTCTGGACAGCAGAGAGGTTGACTTATCAAGAGGATCTCCCTCACTGGAACAAG 240
Db 221 GCTTCCTTCTGGACAGCAGAGAGGTTGACTTATCAAGAGGATCTCCCTCACTGGAACAAG 280
QY 241 CTTAAAGCAGATGAGAAAGTACTTCTCTCATCTCTCAATCTTTAGGCTTTTTCGAGCCAGTGAT 300
Db 281 CTTAAAGCAGATGAGAAAGTACTTCTCTCATCTCTCAATCTTTAGGCTTTTTCGAGCCAGTGAT 340
QY 301 GGAATTGTAATGAAAATTTGGTGAGCGCTTTAGTCAGGAGGTCAGGTTCCAGAGGCT 360
Db 341 GGAATTGTAATGAAAATTTGGTGAGCGCTTTAGTCAGGAGGTCAGGTTCCAGAGGCT 400
QY 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 420
Db 401 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 460
QY 421 CTCATAGACACTTACATCAGA 441
Db 461 CTCATAGACACTTACATCAGA 481

RESULT 12

US-08-307-499-1/c
; Sequence 1, Application US/08307499
; Patent No. 5651972
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; TITLE OF INVENTION: Live Vaccine Vector
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: U.S.A.
; ZIP: 32606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/307,499
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/908,241
FILING DATE: 1-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/908,630
FILING DATE: 29-JUN-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/342,212
FILING DATE: 21-APR-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UF35.1.FWCC1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:

```

; LENGTH: 14176 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3852..4226
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 4585..4887
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5131..5310
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5760..5912
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 6786..7130
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 10148..10513
; US-08-307-499-1

Query Match      40.8%; Score 429.8; DB 1; Length 14176;
Best Local Similarity 66.2%; Pred. No. 3.3e-119; Indels 5; Gaps 2;
Matches 651; Conservative 0; Mismatches 327;

QY 71 AAAGTGAATAAAGTCAATGAAGAGCCACTCTAAGAAAGAGTCTCGCCGGTTGTCA 130
Db 10841 AAAATGATTTTATACAGAAATGGAGCCTATCTTCAAGAGTCTGATCTAGTTCGTTA 10782

QY 131 TCTTTCCAAATCCAGTACCCTGATATTTGGAAAATGATATAAAGGACAGGCTTCTCTTCT 190
Db 10781 TTTTCCCTATTAAAGTATCATGATATCTGAAAAATGTATAAACAATCAGTGGCAAGTTT 10722

QY 191 GGACAGCAGAGAGTGTGACTTATCAAAAGGATCTCCCTACCTGGAACAGCTTAAAGCAG 250
Db 10721 GGACCGTTGAAGAGTAGATTTATCAAAAGATTTAGATGATTTGGGATAAATTAACATAAG 10662

QY 251 ATGAGAAGTACTTCATCTCTACATCTTAGCCCTTTTTCGACGACAGTATGGAATGTAA 310
Db 10661 ACGAAAATACCTTTATAAACAATATACAGATTTTTCGATCTAGTATGATGTATTGTA 10602

QY 311 ATGAAAAATTTGGTGGAGCCCTTTAGTCAGAGGTCGAGGTCAGAGGCTCGCTGTTCT 370
Db 10601 ATGAAATTTAGCGGAAAAGATTTTATGTGATGTACAGTGTTCAGAGGCACGATGTTCT 10542

QY 371 ATGGCTTCAAAATCTCATCGAGATGTCTACACAGATGTACAGTGTTCGCTGATAGACA 430
Db 10541 ATGGATTTCAAAATAGCTATGGAATAATTCATTCAGAAATGATATGATTTATTAATAGATA 10482

QY 431 CTTACATCAGAGATCCCAAGAAAAGGGAATTTTAAATGCAATTTGAAACCATGCCCT 490
Db 10481 CATATGTAAGAGATATAATAGAAAATGTCATTTTAAACCTATAGAAACAATGGAAT 10422

QY 491 ATGTTTAAAGAAAAGCAGATCGGCTTGGCGATGGATAGCAGATAGAAAATCTACTTTTG 550
Db 10421 GCGTAAAAAAGAAAGCTGATTTGGGCCAGAAAATGGATATC---TAGCAACAGGTATATG 10365

QY 551 GGGAAAGAGTGTGGCCCTTTCGCTGTAGAGAGGATTTCTCTCAGGATCTTTTGTG 610
Db 10364 GAGAAAGAGTGTAGCATTTGAGCTGTGGAGGGAATATCTTTTCTGTTCTATTTGCTG 10305

QY 611 CTATATCTCGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTCCAAATGAACCTCA 670
Db 10304 CTATATTTTGGATATAAACAACAGAGATGAUCCCGGATTAACATTTTCTAAATGAACATA 10245

QY 671 TCAGCAGAGATGAAGGACTTCACTGTGATCTTTGCTGCTGATGTTTCCAAATCTTAGTAA 730
Db 10244 TAAGTAGAGACCAAGGTTTACATTTGTGATTTTGGCGTGTAAATGTTTAAACATTTATTAC 10185

```

```

QY 731 ATAAGCCTTCAGAAGAAAGGTCAGGAGATCATTTGTATGCTGTCAAAATTTGAGCAGG 790
Db 10184 ATCCACCATCTAAGGAAGTTATAACGTCGATATCATTTGATGCGGTAAATATAGAAAAGG 10125

QY 791 AGTTTTTAAACAAGACCTTTGCCAGTTGGCCTCATTCGAATGAATTCGATTTTGAAGAAC 850
Db 10124 AGTTTTTGAACAGTTGCTATTTCGGTGGATCTTATAGGTATGAATTTGTTTAAATGTCCTC 10065

QY 851 AGTACATTTGAGTTTGTAGCTGACAGATTACTTGTGAACTTGTGAACTTCTCAAGGTTTTC 910
Db 10064 AGTATATAGAATTCGTCGAGATAGATTATTAACAGAGTTAGGTTGTGAAAAG--TCTCA 10007

QY 911 AGCCAGAAAATCCTTTTGTATTTTATGAAAAACATTTCTTTAGAGAAAAACAAATTTCT 970
Db 10006 ATGTATATAATCCTTTTAGCTTTATGGAGTATATATCACTAGAGGTAGACTAATTTTT 9947

QY 971 TTGAGAAAACAGTTTCAGAGTATACAGCGTTTTCAGCTTATGCGAGAAACACAGATAACG 1030
Db 9946 TCGAAGCAGCAGTTAGTGAATATCAAAAGATGGGGTGTTTACAAATAAAGAGAGAATA 9887

QY 1031 TCTTCACCTTGGATCGAGATTTT 1053
Db 9886 TATTTTCTACGGATATAGATTTT 9864

RESULT 13
US-08-307-499-14
; Sequence 14, Application US/08307499
; Patent No. 5651972
; GENERAL INFORMATION:
; APPLICANT: Meyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: U.S.A.
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/307,499
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,241
; FILING DATE: 1-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,630
; FILING DATE: 29-JUN-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/342,212
; FILING DATE: 21-APR-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF35.1.FWCC1
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14176 base pairs

```

TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS
LOCATION: 138..1460

FEATURE:

NAME/KEY: CDS
LOCATION: 2456..2659

FEATURE:

NAME/KEY: CDS
LOCATION: 2809..3030

FEATURE:

NAME/KEY: CDS
LOCATION: 3070..3330

FEATURE:

NAME/KEY: CDS
LOCATION: 3356..4180

FEATURE:

NAME/KEY: CDS
LOCATION: 4392..5894

FEATURE:

NAME/KEY: CDS
LOCATION: 6171..6398

FEATURE:

NAME/KEY: CDS
LOCATION: 6447..6875

FEATURE:

NAME/KEY: CDS
LOCATION: 6928..7431

FEATURE:

NAME/KEY: CDS
LOCATION: 7454..7858

FEATURE:

NAME/KEY: CDS
LOCATION: 7895..8155

FEATURE:

NAME/KEY: CDS
LOCATION: 8215..8682

FEATURE:

NAME/KEY: CDS
LOCATION: 8715..9539

FEATURE:

NAME/KEY: CDS
LOCATION: 9562..10272

FEATURE:

NAME/KEY: CDS
LOCATION: 10316..11908

FEATURE:

NAME/KEY: CDS
LOCATION: 11971..12780

FEATURE:

NAME/KEY: CDS
LOCATION: 12829..13107

FEATURE:

NAME/KEY: CDS
LOCATION: 13149..14171

FEATURE:

NAME/KEY: CDS
LOCATION: 14171..1499-14

FEATURE:

NAME/KEY: CDS
LOCATION: 1499-14..1500-14

FEATURE:

NAME/KEY: CDS
LOCATION: 908..1000

FEATURE:

NAME/KEY: CDS
LOCATION: 1000..1000

FEATURE:

NAME/KEY: CDS
LOCATION: 1000..1000

FEATURE:

NAME/KEY: CDS
LOCATION: 1000..1000

Query Match 40.8%; Score 429.8; DB 1; Length 14176;

Best Local Similarity 66.2%; Pred. No. 3.3e-119;

Matches 651; Conservative 0; Mismatches 327; Indels 5; Gaps 2;

QY 71 AAAGTGAATAAGTCAATGAAGCCACTCTTAAGAAGAGTTCTCGCCGGTTTGCA 130

DB 3336 AAATGATTTTATACAGATGGAGCCCTATTCTTCAAGAGTCTGATTTAGTTCGTTA 3395

QY 131 TCTTCCCAATCCAGTACCTGATATTGGAATGATAAAGGACAGGCTTCCTCT 190

DB 3396 TTTTCCCTATTAGTATCATGATATCTCGAAATGATAAACAATCAGTGGCAAGTTT 3455

QY 191 GGACAGCAGAAGAGTTGATCTTATCAAGGATCTCCCTCACTGGAAACAAGCAG 250

Db 3456 GGACCGTTGAAGAGTAGATTATCAAAAGATTAGATGTTGGATATAATACTAAAG 3515

QY 251 ATGAGAAGTACTTCACTCTCAATCTTACATCTTTTGGCAGCAGTATGGAATGTA 310

Db 3516 ACGAAAAATACTTTTATAAAAACATATAGCAATTTTTCATCTAGTATGATGTTGTA 3575

QY 311 ATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCAGAGGCTCGCTGTTCT 370

Db 3576 ATGAGAATTTTGGCGAAAGATTATGATGATGATGATGATGATGATGATGATGAT 3635

QY 371 ATGCTTTTCAAAATTTCTCATCGAGATGTTCACTCAGAGATGATCAGTTCGTGATAGACA 430

Db 3636 ATGGAATTTCAATAGCTATGGAATAATTTCACTCAGAAATGATAGTTTATTAATAGATA 3695

QY 431 CTTACATCAGAGATCCAAAGAAAGGGAATTTTATTTAATGCAATTTGAAACCAATGCCCT 490

Db 3696 CATATGTAAGAGATAATATAGAAAAAATGCAATTTATTTAAGCTATAGAAAAATGGAAT 3755

QY 491 ATGTTAAGAAAAAGCAGATTTGGCCTTGGATGGATAGCAGATAGAAAAATCTACTTTTG 550

Db 3756 GCGTAAAAAAGAAAGCTGATTGGGCCAGAAAAATGGATATC---TAGCAACAAGGTATATG 3812

QY 551 GCGAAAGAGTGGTGGCTTTTCTGCTGTAGAAGAGTTCCTCTCAGGATCTTTTGGCTG 610

Db 3813 GAGAAAGAGTAGTAGCATTTTCAGCTGTGAGGGAATATTTCTTTTCGTTCAATTTGCTG 3872

QY 611 CTATATTTCTGGCTAAAGAAAGAGAGTCTTATGCCAGGACTCACATTTTCCAATGAATCA 670

Db 3873 CTATATTTTGGATAAAAAACGAGGATTGATGCCGATTAACATTTTCTAATGAATCA 3932

QY 671 TCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTTCAATGATTTAGTAA 730

Db 3933 TAAAGTAGAGCAGGAGTTTACATTTGATTTTGGCTGTTTAAATGTTTAAACATTTATTAT 3992

QY 731 ATAAGCCTTCAGAAAGAAAGGTCAGGAGATCATTTGATGCTGTCACAAATTCAGCAGG 790

Db 3993 ATCCACCATCTAAGGAAGTTATAACGTCGATAATCATTTGATGCGGTTAATATAGAAAAAG 4052

QY 791 AGTTTTTAACAGAAAGCCTTCCAGTTGGCCTCATTTGGAATGAATTTGATGAAC 850

Db 4053 AGTTTTTGACAGTTGCTATTTCCGTTGGATCTTATAGGTATGAATTTGTTTATGCTC 4112

QY 851 AGTACATGAGTTTGTAGTGAAGATTAATTTGGAATTTGGAATTTGGAATTTGGAATTTTTC 910

Db 4113 AGTATATAGAAATTCGTCGAGATAGATTTAATTAACAGAGTTAGTTGTTGAAAAAG--TCTCA 4170

QY 911 AGGCAGAAATCCTTTTGTATTTTATGAAACATTTTCTTTAGAGGAAACAAATTTCT 970

Db 4171 ATGATATATATCCTTTTGTAGCTTTTATGAGTATATATACATAGAGGTAAGACTAATTTT 4230

QY 971 TTGAGAAACAGTTTTCAGAGTATCAGCGTTTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGT 1030

Db 4231 TCGAACGACGAGTTAGTGAATATCAAAAGATGGGGTGTGTTTACAAATTAAGAGAGAATA 4290

QY 1031 TCTTCACTTGGATGAGATTTT 1053

Db 4291 TATTTTCTACGGATATAGATTTT 4313

RESULT 14

US-09-299-268-1/c

; Sequence 1, Application US/09299268

; Patent No. 6217882

; GENERAL INFORMATION:

; APPLICANT: Moyer, Richard W.

; APPLICANT: Vi uela, Eladio

; APPLICANT: Gibbs, E.P.J.

; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a

; TITLE OF INVENTION: Live Vaccine Vector

; NUMBER OF SEQUENCES: 60

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: David R. Saliwanchik

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: Florida

COUNTRY: U.S.A.

ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/299,268

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/901,127

FILING DATE:

APPLICATION NUMBER: US 07/908,241

FILING DATE: 1-JUL-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/908,630

FILING DATE: 29-JUN-1992

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/342,212

FILING DATE: 21-APR-1992

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Sallwanchik, David R.

REGISTRATION NUMBER: 31,794

REFERENCE/DOCKET NUMBER: UF35.1.FWCC1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 904-375-8100

TELEFAX: 904-372-5800

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 14176 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: unknown

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS

LOCATION: 3852..4226

FEATURE:

NAME/KEY: CDS

LOCATION: 4585..4887

FEATURE:

NAME/KEY: CDS

LOCATION: 5131..5310

FEATURE:

NAME/KEY: CDS

LOCATION: 5760..5912

FEATURE:

NAME/KEY: CDS

LOCATION: 6786..7130

FEATURE:

NAME/KEY: CDS

LOCATION: 10148..10513

US-09-299-268-1

Query Match 40.8%; Score 429.8; DB 3; Length 14176;

Best Local Similarity 66.2%; Pred. No. 3.3e-119;

Matches 651; Conservative 0; Mismatches 327; Indels 5; Gaps 2;

QY 71 AAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTCTCGCCGGTTGTCA 130

Db 10841 AAATGATTTTATACAGAAATGGAGCCCTATTCTTCAAGAGTCTGATTTAGGTTGTTA 10782

QY 131 TCTTCCCAATCCAGTCCCTGATATTGGAAAATGATATAAAGCAGCAGGCTTCCTTCT 190

Db 10781 TTTTCCCTATTAAAGTATCATGATATCTGAAAATGTATAAACAATCAGTGGCAAGTTT 10722

QY 191 GGACAGCAGAAAGAGTTGACTTATCAAAAGATCTCCCTCACTGGAAACAAGCTTAAGCAG 250

Db 10721 GGACCGTTGAAGAGTAGATTATCAAAAGATTAGATGTTGGATATAATTAACATAAG 10662

QY 251 ATGAGAAAGTACTTCTCTCTCACATCTTAGCCCTTTTGGAGCCAGGTATGGAATGTAA 310

Db 10661 ACCAAAATACTTTATATAAAACATATACATAGCTTTTGGCATCTAGTGTGTTGTA 10602

QY 311 ATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAAGTTCACAGGCTCGCTTTCT 370

Db 10601 ATGAGAAATTTAGCGGAAAGATTTTATGTGATGTACAGTGTTCAGAGCCAGATGTTCT 10542

QY 371 ATGGCTTTCAAAATTTCTCATCGAATGTTTCACTCAGAGATGTACAGTTTGTGTATAGACA 430

Db 10541 ATGGATTTCAATAGCTATGGAATAATTTCAATCAGAAATGTATAGTTTATTAATAGATA 10482

QY 431 CTTACATCAGAGATCCCAAGAAAAGGAAATTTTATTAATGCAATTTGAAACCAATGCGCT 490

Db 10481 CATATCTAAGAGATAATATAGAAAAAATGCAATTTATTTAAACGCTATAGAAAAAATGGAAT 10422

QY 491 ATGTTAAGAAAAAGCAGATTTGGCCCTTGCATGGATAGCAGATAGAAAAATCTACTTTTG 550

Db 10421 GCGTAAAAAAGAAAGCTGATTTGGGCCAGAAAAATGGATATC---TAGCAACAAGGTATATG 10365

QY 551 GGGAAAGAGTGTGGCTTTTGTCTGTAGAAAGAGTCTTCTCTCAGGATCTTTTGTCTG 610

Db 10364 GAGAAAGAGTAGTAGCATTTGCAGCTGTGAGGGAATATTTCTTTCTGTTTCAATTTGCTG 10305

QY 611 CTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGAGTCACTTTTCCAAATCAACTCA 670

Db 10304 CTATATTTTGGATAAAAAAGCAGAGGATGATGCCGATTAACATTTTCTTAATGAACATA 10245

QY 671 TCAGCAGAGATGAAGACTTCACTGTGCTTCTGCTTCCCTGCTGATGTTCCCAATCACTAGTAA 730

Db 10244 TAAGTAGACGCAAGGTTTACATTTGTGATTTTGGCTGTTTAATGTTTAAACATTTATTTAC 10185

QY 731 ATAAGCCTTCAGAAAGAAAGGTCAGGAGAGATCATTTGTGATGCTGTCAAAATTCAGCAGG 790

Db 10184 ATCCACCATCTAAGGAAGTTATAACGTCGATAATCATTTGATCGGTTAATATAGAAAAGG 10125

QY 791 AGTTTTTAACAGAAAGCCTTGGCCAGTTGGCCTCAATTTGGAATGAATTTGATTTGATGAAC 850

Db 10124 AGTTTTTGACAGTTGCTATTTCCGGTGGATCTTATAGGTATGAATTTGTTTAAATGCTC 10065

QY 851 AGTACATTGAGTTTGTAGCTGCAGAGATTAATTTGGAATGAATTTGATTTGATTTTCT 910

Db 10064 AGTATATAGAATTCGTCGAGATAGATTATTAACAGAGTTAGGTTGTGAAAAG--TCTCA 10007

QY 911 AGGCAGAAAATCCTTTTGTATTTTATGAAAAACATTTCTTTAGAAAGAAAAACAAATTTCT 970

Db 10006 ATGTATATATCTTTTGTAGCTTTTATGGAGTATATATCACTAGAAAGTTAGACTAATTTT 9947

QY 971 TTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGGTTTTCAGTTATGGCAGAAACCAAGATAACG 1030

Db 9946 TCGAACGACGAGTTAGTGAATATCAAAAGATGGGGTGTTTTACAAATAAAGAGAGAATA 9887

QY 1031 TCTTCACTTGGATGCAGATTTT 1053

Db 9886 TATTTTCTACGGATATAGATTTT 9864

RESULT 15

US-09-299-268-14

; Sequence 14, Application US/09299268

; Patent No. 6217882

; GENERAL INFORMATION:

; APPLICANT: Moyer, Richard W.

; APPLICANT: Vi uela, Eladio

; APPLICANT: Gibbs, E.P.J.

; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a

; TITLE OF INVENTION: Live Vaccine Vector

; NUMBER OF SEQUENCES: 60

; CORRESPONDENCE ADDRESS:

ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: U.S.A.
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/299,268
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/901,127
FILING DATE:
APPLICATION NUMBER: US 07/908,241
FILING DATE: 1-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/908,630
FILING DATE: 29-JUN-1992
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/342,212
FILING DATE: 21-APR-1992
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UF35.1.FWCC1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 14176 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 138...1460
FEATURE:
NAME/KEY: CDS
LOCATION: 2456...2659
FEATURE:
NAME/KEY: CDS
LOCATION: 2809...3030
FEATURE:
NAME/KEY: CDS
LOCATION: 3070...3330
FEATURE:
NAME/KEY: CDS
LOCATION: 3356...4180
FEATURE:
NAME/KEY: CDS
LOCATION: 4392...5894
FEATURE:
NAME/KEY: CDS
LOCATION: 6171...6398
FEATURE:
NAME/KEY: CDS
LOCATION: 6447...6875
FEATURE:
NAME/KEY: CDS
LOCATION: 6928...7431
FEATURE:
NAME/KEY: CDS
LOCATION: 7454...7858
FEATURE:

NAME/KEY: CDS
LOCATION: 7895...8155
FEATURE:
NAME/KEY: CDS
LOCATION: 8215...8682
FEATURE:
NAME/KEY: CDS
LOCATION: 8715...9539
FEATURE:
NAME/KEY: CDS
LOCATION: 9562...10272
FEATURE:
NAME/KEY: CDS
LOCATION: 10316...11908
FEATURE:
NAME/KEY: CDS
LOCATION: 11971...12780
FEATURE:
NAME/KEY: CDS
LOCATION: 12829...13107
FEATURE:
NAME/KEY: CDS
LOCATION: 13149...14171
US-09-299-268-14
Query Match 40.8%; Score 429.8; DB 3; Length 14176;
Best Local Similarity 66.2%; Pred. No. 3.3e-119;
Matches 651; Conservative 0; Mismatches 327; Indels 5; Gaps 2;
QY 71. AAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTCAAGAAAGAGTTCTCGCGGTTTGTCA 130
DB 3336 AAATGATTTTATACAGATCGAGCCTATTCTTCAAGAGTCTGATCTTAGGTTCTGTTA 3395
QY 131 TCTTTCCAAATCCAGTACCCCTGATATTTGGAAATATGTAACAGGACAGGCTTCCTTCT 190
DB 3396 TTTTCCCTATTAAAGTATCATGATCTCGAAATGTAATAACAATCAGTGGCAAGTTT 3455
QY 191 GGACAGCAGAGAGTTGACTTATCAAGAGATCTCCCTCACTGGAAACAAGCTTAAGCAG 250
DB 3456 GGACCGTTGAAGAGTAGATTATCAAAAGATTAGATGTTGGGATAAATAACTAAG 3515
QY 251 ATGAGAAAGTACTTCTCATCTCTCACATCTTAGCCCTTTTTCAGCCAGGATGTAATGTA 310
DB 3516 ACGAAATACTTTTATAAACAATATAGTATCTAGTATCTAGTATGTTGTA 3575
QY 311 ATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCCAGAGGCTCGCTTCT 370
DB 3576 ATGAGAAATTTAGCGGAAAGATTTTATGTGGATGTACAGTGTTCAGAGGACGATGTTCT 3635
QY 371 ATGGCTTTCAAATCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTGTGTATAGACA 430
DB 3636 ATGGATTTCAAATAGCTATGGAATAATTCATTTCAAGAAATGTATAGTTTATTAAGATA 3695
QY 431 CTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGAATGAAACCATGCCCT 490
DB 3696 CATATGTAGAGATAATATAGAAATAATGCATTTATTAACGCTATAGAAACAATGGAAT 3755
QY 491 ATGTTAAGAAAAAGCAGATTGGGCTTGGATGATAGCAGATAGAAAAATCTACTTTTG 550
DB 3756 GCGTAAAAAAGAAAGCTGATTGGCCAGAAAAATGGATATC--TAGCAACAAGTATATG 3812
QY 551 GGGAAAGAGTGGGCTTTGCTGTGTAGAAAGAGTTTCTCTCAGGATCTTTTGGCTG 610
DB 3813 GAGAAAGAGTAGTAGCATTTGCAGCTGTGGAGGAAATTTCTTTCTGTTTCTTCTGCTG 3872
QY 611 CTATATTTCTGGCTAAAGAGAGAGAGGCTTATGTCAGGACTCACTTTTCCAAATCACTCA 670
DB 3873 CTATATTTTGGATAAAAAACAAGAGATGATGCCCGGATTAACATTTCTTAATGAACATA 3932
QY 671 TCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTCTGCTGATGTTTCCAAATCTAGTAA 730
DB 3933 TAAGTAGACGACGAGGTTTACATTTGTGATTTTGGTGTAAATGTTTAAACATTTATAC 3992

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 30, 2005, 01:51:37 ; Search time 676.5 Seconds
(without alignments)
12855.457 Million cell updates/sec

Title: US-10-698-228-12
Perfect score: 1053
Sequence: 1 atggcgaccggaagcc.....tcacctggatgcagatttt 1053

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9784742 seqs, 4129495052 residues

Total number of hits satisfying chosen parameters: 19569484

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

```

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
17: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:
18: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:
19: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:
20: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:
21: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:
22: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:
23: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
24: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
25: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:
26: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:
27: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:
28: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1053	100.0	1053	19	US-10-698-228-12
2	1053	100.0	1081	19	US-10-698-228-4
3	1051.4	99.8	1053	19	US-10-698-228-2
4	1051.4	99.8	4955	18	US-10-172-118-71
5	1051.4	99.8	4955	19	US-10-342-887-71

6	1051.4	99.8	4955	19	US-10-698-228-3
7	580	55.1	2482	24	US-10-220-335-514
8	580	55.1	2500	9	US-09-954-456-724
9	580	55.1	2500	9	US-09-954-456-1169
10	580	55.1	2500	9	US-09-954-456-1827
11	580	55.1	2500	19	US-10-641-643-1370
12	580	55.1	2500	21	US-10-733-878-458
13	580	55.1	2500	22	US-10-843-641A-3751
14	580	55.1	2500	22	US-10-843-641A-4196
15	580	55.1	2500	22	US-10-843-641A-4854
16	580	55.1	2500	24	US-10-756-149-713
17	578.4	54.9	1989	9	US-09-925-301-505
18	578.4	54.9	2216	16	US-10-084-817-342
19	570.4	54.2	2113	22	US-10-764-420-1636
20	569	54.0	2641	24	US-10-220-335-170
21	562.6	53.4	1328	19	US-10-403-571-75
22	545.4	51.8	977	18	US-10-264-237-790
23	514	48.8	1371	24	US-10-450-763-15278
24	463	44.0	1289	26	US-11-097-143-19181
25	447.8	42.5	186854	22	US-10-872-156-34
26	442	42.0	1146	9	US-09-822-830A-174
27	433.8	41.2	1218	16	US-10-128-714-7245
28	420.4	39.9	3945	26	US-11-097-143-19180
29	404.8	38.4	963	18	US-10-369-493-34052
30	398.6	37.9	1146	16	US-10-128-714-2245
31	394.2	37.4	1173	18	US-10-369-493-46252
32	393.6	37.4	1242	17	US-10-032-585-6537
33	388.2	36.9	1292	18	US-10-369-493-27809
34	382	36.3	1200	18	US-10-369-493-45857
35	380.6	36.1	1248	19	US-10-424-599-126843
36	372.8	35.4	1361	19	US-10-424-599-40471
37	365.8	34.7	1586	19	US-10-424-599-73024
38	363.2	34.5	957	21	US-10-653-047-43
39	363.2	34.5	1206	18	US-10-320-797-2145
40	350.6	33.3	1314	16	US-10-128-714-1245
41	350.6	33.3	1450	16	US-10-128-714-6245
42	350.6	33.3	3314	16	US-10-128-714-2245
43	350.6	33.3	3450	16	US-10-128-714-5245
44	349.8	33.2	629	24	US-10-450-763-6851
45	328.2	31.2	560	9	US-09-864-761-12381

ALIGNMENTS

RESULT 1
US-10-698-228-12
; Sequence 12, Application US/10698228
; Publication No. US20040072253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 12
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-12

Query Match 100.0%; Score 1053; DB 19; Length 1053;

Best Local Similarity 100.0%; Pred. No. 9.3e-295;			
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	ATGGCGGACCCGGAAGCCGGAAGCGCGCGGTGGATCAGGATGAGAGATCATCTTCA	60
Db	1	ATGGCGGACCCGGAAGCCGGAAGCGCGCGGTGGATCAGGATGAGAGATCATCTTCA	60
QY	61	GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTAAGAAAGAGTTCTCG	120
Db	61	GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTAAGAAAGAGTTCTCG	120
QY	121	CGGTTTGTCATCTTTCCCAATCAGTACCTCGATATTTGGAAATGTATAAACAGGCACAG	180
Db	121	CGGTTTGTCATCTTTCCCAATCAGTACCTCGATATTTGGAAATGTATAAACAGGCACAG	180
QY	181	GCTTCCTTCTGACAGCAGAGAGGTGACTTATCAAGAGATCTCCCTCACTGGAAACAAG	240
Db	181	GCTTCCTTCTGACAGCAGAGAGGTGACTTATCAAGAGATCTCCCTCACTGGAAACAAG	240
QY	241	CTTAAGCAGATGAGAAGTACTTCACTCTCAATCTTAGCCCTTTTTCAGCCAGTGAT	300
Db	241	CTTAAGCAGATGAGAAGTACTTCACTCTCAATCTTAGCCCTTTTTCAGCCAGTGAT	300
QY	301	GGAAATGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT	360
Db	301	GGAAATGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT	360
QY	361	CGCTGTTTCTATGGCTTTCAAAATCTCATCGAAGTGTTCATCAGAGATGACAGTTTG	420
Db	361	CGCTGTTTCTATGGCTTTCAAAATCTCATCGAAGTGTTCATCAGAGATGACAGTTTG	420
QY	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA	480
Db	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA	480
QY	481	ACCATGCGCTATGTTAAGAAAGAGCAGATTGGGCTTCGCGATGATAGCAGATAGAAA	540
Db	481	ACCATGCGCTATGTTAAGAAAGAGCAGATTGGGCTTCGCGATGATAGCAGATAGAAA	540
QY	541	TCTATCTTTGGGAAAGAGTGTGGCTTTGCTGTGAAGAGGATTTTCTTCAGGA	600
Db	541	TCTATCTTTGGGAAAGAGTGTGGCTTTGCTGTGAAGAGGATTTTCTTCAGGA	600
QY	601	TCTTTTGTGCTATATTCTGGCTAAGAGAGAGTCTTATGCCAGGACTCACTTTTCC	660
Db	601	TCTTTTGTGCTATATTCTGGCTAAGAGAGAGTCTTATGCCAGGACTCACTTTTCC	660
QY	661	AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGGCTGTATGTTCCAA	720
Db	661	AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGGCTGTATGTTCCAA	720
QY	721	TACTTAGTAAATTAAGCCTTCAAGAAAGGGTCAAGGAGATCATTTGATGCTGTCAAA	780
Db	721	TACTTAGTAAATTAAGCCTTCAAGAAAGGGTCAAGGAGATCATTTGATGCTGTCAAA	780
QY	781	ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTTGGCTCATTTGGAAATGAATTGCATT	840
Db	781	ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTTGGCTCATTTGGAAATGAATTGCATT	840
QY	841	TTGATGAACACGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACTTGGATTCTCA	900
Db	841	TTGATGAACACGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACTTGGATTCTCA	900
QY	901	AAGGTTTTTTCAGGCGAAGAAATCCTTTTGATTTTATGGAAGAAACATTTCTTTAGAAGAAA	960
Db	901	AAGGTTTTTTCAGGCGAAGAAATCCTTTTGATTTTATGGAAGAAACATTTCTTTAGAAGAAA	960
QY	961	ACAAATTTCTTTCAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC	1020
Db	961	ACAAATTTCTTTCAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC	1020
QY	1021	ACAGATAACGCTCTTACCTTGGATGCAGATTTT 1053	

Db	1021	ACAGATAACGCTCTTACCTTGGATGCAGATTTT 1053	
RESULT 2			
US-10-698-228-4			
; Sequence 4, Application US/10698228			
; Publication No. US2004007253A1			
; GENERAL INFORMATION:			
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA			
; TITLE OF INVENTION: New Protein and its DNA			
; FILE REFERENCE: 2619WO0P			
; CURRENT APPLICATION NUMBER: US/10/698,228			
; CURRENT FILING DATE: 2003-10-30			
; PRIOR APPLICATION NUMBER: US/10/019,733			
; PRIOR FILING DATE: 2001-12-28			
; PRIOR APPLICATION NUMBER: JP 11-181131			
; PRIOR FILING DATE: 1999-06-28			
; PRIOR APPLICATION NUMBER: JP 11-192391			
; PRIOR FILING DATE: 1999-07-06			
; PRIOR APPLICATION NUMBER: JP 2000-017770			
; PRIOR FILING DATE: 2000-01-21			
; NUMBER OF SEQ ID NOS: 14			
; SEQ ID NO 4			
; LENGTH: 1081			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; OTHER INFORMATION:			
US-10-698-228-4			
Query Match 100.0%; Score 1053; DB 19; Length 1081;			
Best Local Similarity 100.0%; Pred. No. 9.4e-295;			
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	ATGGCGGACCCGGAAGCCGGAAGCGCGCGGTGGATCAGGATGAGAGATCATCTTCA	60
Db	20	ATGGCGGACCCGGAAGCCGGAAGCGCGCGGTGGATCAGGATGAGAGATCATCTTCA	79
QY	61	GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCG	120
Db	80	GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCG	139
QY	121	CGGTTTGTCATCTTTCCCAATCCAGTACCTCGATATTTGGAAATGTATAAACAGGCACAG	180
Db	140	CGGTTTGTCATCTTTCCCAATCCAGTACCTCGATATTTGGAAATGTATAAACAGGCACAG	199
QY	181	GCTTCCTTCTGACAGCAGAGAGGTGACTTATCAAGAGATCTCCCTCACTGGAAACAAG	240
Db	200	GCTTCCTTCTGACAGCAGAGAGGTGACTTATCAAGAGATCTCCCTCACTGGAAACAAG	259
QY	241	CTTAAGCAGATGAGAAGTACTTCACTCTCATCTTAGCCCTTTTTCAGCCAGTGAT	300
Db	260	CTTAAGCAGATGAGAAGTACTTCACTCTCATCTTAGCCCTTTTTCAGCCAGTGAT	319
QY	301	GGAAATGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCT	360
Db	320	GGAAATGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCT	379
QY	361	CGCTGTTTCTATGGCTTTTCAAAATTTCTCATCGAATGTTTATTAATGCAATGAA	420
Db	380	CGCTGTTTCTATGGCTTTTCAAAATTTCTCATCGAATGTTTATTAATGCAATGAA	439
QY	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA	480
Db	440	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA	499
QY	481	ACCATGCGCTATGTTAAGAAAGAGCAGATTGGGCTTCGCGATGATAGCAGATAGAAA	540
Db	500	ACCATGCGCTATGTTAAGAAAGAGCAGATTGGGCTTCGCGATGATAGCAGATAGAAA	559
QY	541	TCTACTTTTGGGAAAGAGTGTGGCTTTTGTGCTGTAGAAGAGTTTTCTTCTCAGGA	600
Db	560	TCTACTTTTGGGAAAGAGTGTGGCTTTTGTGCTGTAGAAGAGTTTTCTTCTCAGGA	619

```
QY 601 TCTTTTGCTGCTATATTCCTGCTAAAGAGAGAGAGGCTTATGCGCAGGACTCATCTTTTCC 660
Db 620 TCTTTTGCTGCTATATTCCTGCTAAAGAGAGAGGCTTATGCGCAGGACTCATCTTTTCC 679
QY 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGGCTGATGTTCCAA 720
Db 680 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGGCTGATGTTCCAA 739
QY 721 TACTTAGTAATAAAGCCTTCAGAAAGAGGCTCAGGAGATCATTTGATGCTGTCAAA 780
Db 740 TACTTAGTAATAAAGCCTTCAGAAAGAGGCTCAGGAGATCATTTGATGCTGTCAAA 799
QY 781 ATTGAGCAGGAGTTTAAACAGAAAGCCTTGCCAGTTGGCTCATTTGGAATGAATTCATT 840
Db 800 ATTGAGCAGGAGTTTAAACAGAAAGCCTTGCCAGTTGGCTCATTTGGAATGAATTCATT 859
QY 841 TTGATGAACAGTACATCAGTTGTTAGCTGACAGATTAATTTGTGGAACCTTGGATTTCTCA 900
Db 860 TTGATGAACAGTACATCAGTTGTTAGCTGACAGATTAATTTGTGGAACCTTGGATTTCTCA 919
QY 901 AAGGTTTTTCAGGCAGAAATCCTTTTGATTTTATGGAACCAATTTCTTTAGAGGAAAA 960
Db 920 AAGGTTTTTCAGGCAGAAATCCTTTTGATTTTATGGAACCAATTTCTTTAGAGGAAAA 979
QY 961 ACAAAATTTCTTGAGAAACGAGTTTCAGAGATATCAGCGTTTTCAGTTATGCGAGAAACC 1020
Db 980 ACAAAATTTCTTGAGAAACGAGTTTCAGAGATATCAGCGTTTTCAGTTATGCGAGAAACC 1039
QY 1021 ACAGATAACGCTCTTCACTTGGATGCGAGATTTT 1053
Db 1040 ACAGATAACGCTCTTCACTTGGATGCGAGATTTT 1072
```

RESULT 3

```
US-10-698-228-2
; Sequence 2, Application US/10698228
; Publication No. US2004007253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-2
```

```
Query Match 99.8%; Score 1051.4; DB 19; Length 1053;
Best Local Similarity 99.9%; Pred. No. 2.7e-294;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 ATGGCGGACCCGGAAGGCGGAGCGGCTGGATCAGGATCAGAGATCATCTTCA 60
Db 1 ATGGCGGACCCGGAAGGCGGAGCGGCTGGATCAGGATCAGAGATCATCTTCA 60
QY 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGGCCACTCTTAAGAAAGAGTTCTCGC 120
```

```
QY 121 CGGTTTGTCTATCTTTTCCAATCCAGTACCTGATATTTGGAAAAATGATAAAACAGGCACAG 180
Db 121 CGGTTTGTCTATCTTTTCCAATCCAGTACCTGATATTTGGAAAAATGATAAAACAGGCACAG 180
QY 181 GCTTCTTCTTGGACAGCAGAAAGAGTTGACTTATCAAAGGATCTCCCTCACTGGAAACAAG 240
Db 181 GCTTCTTCTTGGACAGCAGAAAGAGTTGACTTATCAAAGGATCTCCCTCACTGGAAACAAG 240
QY 241 CTTAAAGCAGATCAGAGTACTTCTCATCTCTCACATCTTAGCCTTTTGGAGCCAGTGTAT 300
Db 241 CTTAAAGCAGATCAGAGTACTTCTCATCTCTCACATCTTAGCCTTTTGGAGCCAGTGTAT 300
QY 301 GGAATTTGTAATCAAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGTTCCAGAGGCT 360
Db 301 GGAATTTGTAATCAAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGTTCCAGAGGCT 360
QY 361 CGCTGTTTCTATCGGCTTTCAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATCGGCTTTCAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
QY 421 CTGATAGACACTTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAA 480
Db 421 CTGATAGACACTTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAA 480
QY 481 ACCATGCCCTTATCTTAAGAAAAAGCAGATTGGGCTTTGGATGATGATAGAGATAGAAAA 540
Db 481 ACCATGCCCTTATCTTAAGAAAAAGCAGATTGGGCTTTGGATGATGATAGAGATAGAAAA 540
QY 541 TCTACTTTTGGGAAAAAGAGTGGTGGCTTTGCTGCTGTAGAAAGAGATTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAAAGAGTGGTGGCTTTGCTGCTGTAGAAAGAGATTTTCTTCTCAGGA 600
QY 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGTCTTATGCGCAGGACTCACATTTTTC 660
Db 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGTCTTATGCGCAGGACTCACATTTTTC 660
QY 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTCCCTGATGTTCCAA 720
Db 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTCCCTGATGTTCCAA 720
QY 721 TACTTAGTAATAAAGCCTTCAGAAAGAGGCTCAGGAGATCATTTGTTGATGCTGTCAAA 780
Db 721 TACTTAGTAATAAAGCCTTCAGAAAGAGGCTCAGGAGATCATTTGTTGATGCTGTCAAA 780
QY 781 ATTGAGCAGGAGTTTAAACAGAAAGCCTTGCCAGTTGGCTTCAATGGAATGAATTCGATT 840
Db 781 ATTGAGCAGGAGTTTAAACAGAAAGCCTTGCCAGTTGGCTTCAATGGAATGAATTCGATT 840
QY 841 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTAATTTGTTGGAACCTTGGATTTCTCA 900
Db 841 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTAATTTGTTGGAACCTTGGATTTCTCA 900
QY 901 AAGGTTTTTTCAGGCAGAAATCCTTTTGATTTTATGGAACCAATTTCTTTAGAAAGAAAA 960
Db 901 AAGGTTTTTTCAGGCAGAAATCCTTTTGATTTTATGGAACCAATTTCTTTAGAAAGAAAA 960
QY 961 ACAAAATTTCTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC 1020
Db 961 ACAAAATTTCTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC 1020
QY 1021 ACAGATAACGCTCTTCACTTGGATGCGAGATTTT 1053
Db 1021 ACAGATAACGCTCTTCACTTGGATGCGAGATTTT 1053
```

RESULT 4

```
US-10-172-118-71
; Sequence 71, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
```

APPLICANT: Mao, Mao
APPLICANT: Roberts, Chris
APPLICANT: Van 't Veer, Laura
APPLICANT: Van de Vijver, Marc
APPLICANT: Bernards, Rene
TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
FILE REFERENCE: 9301-175-999
CURRENT APPLICATION NUMBER: US/10/172,118
CURRENT FILING DATE: 2002-06-14
PRIOR APPLICATION NUMBER: 60/380,770
PRIOR FILING DATE: 2002-05-14
NUMBER OF SEQ ID NOS: 2699
SEQ ID NO 71
LENGTH: 4955
TYPE: DNA
ORGANISM: Homo sapiens
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: AB036063
DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-71

Query Match 99.8%; Score 1051.4; DB 18; Length 4955;
Best Local Similarity 99.9%; Pred. No. 6.8e-294;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGGCGACCCGGAAGCGCGGAGCGCGCGCTGGATCAGGATGAGAGATCATCTTCA 60
245 ATGGCGACCCGGAAGCGCGGAGCGCGCGCTGGATCAGGATGAGAGATCATCTTCA 304

61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
305 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 364

121 CGGTTTGTCACTTTTCCAATCAGTACCTTGATTTTGGAAAATGTATAAACAGGCACAG 180
365 CGGTTTGTCACTTTTCCAATCAGTACCTTGATTTTGGAAAATGTATAAACAGGCACAG 424

181 GCTTCCTTCTGGACAGCAGAGAGTTGACTTATCAAGAGATCTCCCTCACTGGAAACAG 240
425 GCTTCCTTCTGGACAGCAGAGAGTTGACTTATCAAGAGATCTCCCTCACTGGAAACAG 484

241 CTTAAGCAGATGAGAAGTACTTCACTCTCACTTTAGCGCTTTTTCAGCCAGTGAT 300
485 CTTAAGCAGATGAGAAGTACTTCACTCTCACTTTAGCGCTTTTTCAGCCAGTGAT 544

301 GGAATTTGTAATGAAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGT 360
545 GGAATTTGTAATGAAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGT 604

361 CGCTGTTTCTATGGCTTTCAAAATTCATCTCGAATGTTTCACTCAGAGATGACAGTTTG 420
605 CGCTGTTTCTATGGCTTTCAAAATTCATCTCGAATGTTTCACTCAGAGATGACAGTTTG 664

421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTAAATGCAATGAA 480
665 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTAAATGCAATGAA 724

481 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGGCTTTGCGATGATAGCAGATAGAAA 540
725 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGGCTTTGCGATGATAGCAGATAGAAA 784

541 TCTACTTTTGGGAAAGAGTGTGGCTTTTGTCTGTAGAGAGAGTTTCTTCTCAGGA 600
785 TCTACTTTTGGGAAAGAGTGTGGCTTTTGTCTGTAGAGAGAGTTTCTTCTCAGGA 844

601 TCTTTTGTCTATATCTCGCTAAGAGAGAGTCTTATGCCAGGACTCACTTTTCC 660
845 TCTTTTGTCTATATCTCGCTAAGAGAGAGTCTTATGCCAGGACTCACTTTTCC 904

661 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTTGTCTGCTGATGTTCCAA 720
905 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTTGTCTGCTGATGTTCCAA 964

1052 TACTTAGTAAATAAGCCTTCAGAGAAAGGGTTCAGGAGATCATTTGTTGATGCTGTCAA 780
1024 TACTTAGTAAATAAGCCTTCAGAGAAAGGGTTCAGGAGATCATTTGTTGATGCTGTCAA 1024
840 ATTGAGCAGGAGTCTTTTAAACAGAGCCTTCCAGTTGGCCTCATTCGGAATGAATTCATT 840
1084 ATTGAGCAGGAGTCTTTTAAACAGAGCCTTCCAGTTGGCCTCATTCGGAATGAATTCATT 1084
900 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTTACTTGTGGAACCTTGGATTTCA 900
1144 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTTACTTGTGGAACCTTGGATTTCA 1144
960 AAGGTTTTTTCAGGCGAGAAATCCTTTTGTATTTTATGAAAAACATTTCTTTAGAGGAAA 960
1204 AAGGTTTTTTCAGGCGAGAAATCCTTTTGTATTTTATGAAAAACATTTCTTTAGAGGAAA 1204
1020 ACAAATTTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC 1020
1264 ACAAATTTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC 1264
1053 ACAGATAACGCTCTTCACCTTGGATGCGAGATTTT 1053
1297 ACAGATAACGCTCTTCACCTTGGATGCGAGATTTT 1297

RESULT 5
US-10-342-887-71
; Sequence 71, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; PRIOR FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 71
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-71

Query Match 99.8%; Score 1051.4; DB 19; Length 4955;
Best Local Similarity 99.9%; Pred. No. 6.8e-294;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGGCGACCCGGAAGCGCGGAGCGCGCGCTGGATCAGGATGAGAGATCATCTTCA 60
245 ATGGCGACCCGGAAGCGCGGAGCGCGCGCTGGATCAGGATGAGAGATCATCTTCA 304

61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
305 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 364

121 CGGTTTGTCACTTTTCCAATCAGTACCTTGATTTTGGAAAATGTATAAACAGGCACAG 180
365 CGGTTTGTCACTTTTCCAATCAGTACCTTGATTTTGGAAAATGTATAAACAGGCACAG 424

181 GCTTCCTTCTGGACAGCAGAGAGTTGACTTATCAAGAGATCTCCCTCACTGGAAACAG 240
425 GCTTCCTTCTGGACAGCAGAGAGTTGACTTATCAAGAGATCTCCCTCACTGGAAACAG 484

241 CTTAAGCAGATGAGAAGTACTTCACTCTCACTTTAGCGCTTTTTCAGCCAGTGAT 300
485 CTTAAGCAGATGAGAAGTACTTCACTCTCACTTTAGCGCTTTTTCAGCCAGTGAT 544

301 GGAATTTGTAATGAAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGT 360
545 GGAATTTGTAATGAAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGT 604

361 CGCTGTTTCTATGGCTTTCAAAATTCATCTCGAATGTTTCACTCAGAGATGACAGTTTG 420
605 CGCTGTTTCTATGGCTTTCAAAATTCATCTCGAATGTTTCACTCAGAGATGACAGTTTG 664

421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTAAATGCAATGAA 480
665 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTAAATGCAATGAA 724

481 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGGCTTTGCGATGATAGCAGATAGAAA 540
725 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGGCTTTGCGATGATAGCAGATAGAAA 784

541 TCTACTTTTGGGAAAGAGTGTGGCTTTTGTCTGTAGAGAGAGTTTCTTCTCAGGA 600
785 TCTACTTTTGGGAAAGAGTGTGGCTTTTGTCTGTAGAGAGAGTTTCTTCTCAGGA 844

601 TCTTTTGTCTATATCTCGCTAAGAGAGAGTCTTATGCCAGGACTCACTTTTCC 660
845 TCTTTTGTCTATATCTCGCTAAGAGAGAGTCTTATGCCAGGACTCACTTTTCC 904

661 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTTGTCTGCTGATGTTCCAA 720
905 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTTGTCTGCTGATGTTCCAA 964

Db 425 GCTTCCTCTCGGACACGACGAGAGGTCGACTTATCAAAAGGATCTCCCTCACTCGAAACAAG 484
QY 241 CTTAAACGACAGATGAGAACTACTCTCTCATCTCTAGCCCTTTTGGCAGCCAGTGAT 300
Db 485 CTTAAACGACAGATGAGAACTACTCTCTCATCTCTAGCCCTTTTGGCAGCCAGTGAT 544
QY 301 GGAATTGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCGAGGTTCCAGAGGCT 360
Db 545 GGAATTGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCGAGGTTCCAGAGGCT 604
QY 361 CGCTGTTCTATGGCTTCAAAATCTCATCGAGAAAGTTTCACTCAGAGATGTACAGTTTG 420
Db 605 CGCTGTTCTATGGCTTCAAAATCTCATCGAGAAAGTTTCACTCAGAGATGTACAGTTTG 664
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 480
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 724
QY 481 ACCATGCCCTATGTTAAAGAAAGGAGATTTGGCCCTTGGATGGATAGCAGATAGAAAA 540
Db 725 ACCATGCCCTATGTTAAAGAAAGGAGATTTGGCCCTTGGATGGATAGCAGATAGAAAA 784
QY 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAGAGGATTTTCTCTCAGGA 600
Db 785 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAGAGGATTTTCTCTCAGGA 844
QY 601 TCTTTTGTCTGTATATCTCGGCTTAAAGAAAGAGAGTCTTATGCGAGACTCACTTTTTC 660
Db 845 TCTTTTGTCTGTATATCTCGGCTTAAAGAAAGAGAGTCTTATGCGAGACTCACTTTTTC 904
QY 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 905 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 964
QY 721 TACTTAGTAAATGAAGCTTTTAAAGAAAGGCTCAGAGAGATCATTTGATGCTGCAAA 780
Db 965 TACTTAGTAAATGAAGCTTTTAAAGAAAGGCTCAGAGAGATCATTTGATGCTGCAAA 1024
QY 781 ATTGACGAGGATTTTAAAGAAAGGCTTGCAGAGTGGCTCATTTGGAATGAATTGCATT 840
Db 1025 ATTGACGAGGATTTTAAAGAAAGGCTTGCAGAGTGGCTCATTTGGAATGAATTGCATT 1084
QY 841 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTTACTTGTGAACTTGGATTTCTCA 900
Db 1085 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTTACTTGTGAACTTGGATTTCTCA 1144
QY 901 AAGGTTTTTCAGGCAAGAAATCCTTTTGAATTTTATGGAACAAATTTCTTTAGAGGAAAA 960
Db 1145 AAGGTTTTTCAGGCAAGAAATCCTTTTGAATTTTATGGAACAAATTTCTTTAGAGGAAAA 1204
QY 961 ACAAAATTTCTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020
Db 1205 ACAAAATTTCTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1264
QY 1021 ACAGATAAGCTCTTCACTTGGATGCAAGATTTT 1053
Db 1265 ACAGATAAGCTCTTCACTTGGATGCAAGATTTT 1297

RESULT 6

US-10-698-228-3
; Sequence 3, Application US/10698228
; Publication No. US2004007253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28

; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 3
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-3

Query Match 99.8% Score 1051.4; DB 19; Length 4955;
Best Local Similarity 99.9%; Pred. No. 6.8e-294;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAAGCGCGGAGCGGCTGGATCAGGATCAGGATCATCTTCA 60
Db 245 ATGGGCGACCCGGAAAGCGCGGAGCGGCTGGATCAGGATCAGGATCATCTTCA 304
QY 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 305 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 364
QY 121 CGCTTTGTCTATCTTCCAAATCCAGTACCTCTGATATTTGGAAAAATGTATAACAGGCACAG 180
Db 365 CGCTTTGTCTATCTTCCAAATCCAGTACCTCTGATATTTGGAAAAATGTATAACAGGCACAG 424
QY 181 GCCTTCTTCTGGACAGCAGAGAGGTTGACTTATCAAAAGGATCTCCCTCACTGGAAACAAG 240
Db 425 GCCTTCTTCTGGACAGCAGAGAGGTTGACTTATCAAAAGGATCTCCCTCACTGGAAACAAG 484
QY 241 CTTAAAGCAGATCAGAAAGTACTTCTCATCTCTCACTCTTTAGCCTTTTTCGAGCCAGTGAT 300
Db 485 CTTAAAGCAGATCAGAAAGTACTTCTCATCTCTCACTCTTTAGCCTTTTTCGAGCCAGTGAT 544
QY 301 GGAATTGTAAATGAAATTTTGGTGGAGCGCTTTAGTCAGGAGTGCGAGGTTCCAGAGGCT 360
Db 545 GGAATTGTAAATGAAATTTTGGTGGAGCGCTTTAGTCAGGAGTGCGAGGTTCCAGAGGCT 604
QY 361 CGCTGTTCTTATCGCTTTTCAAAATTTCTCATCGAGAAATGTTTCACTCAGAGATGTACAGTTTG 420
Db 605 CGCTGTTCTTATCGCTTTTCAAAATTTCTCATCGAGAAATGTTTCACTCAGAGATGTACAGTTTG 664
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 724
QY 481 ACCATGCCCTATGTTAAAGAAAGGAGATTTGGCCCTTGGATGGATAGCAGATAGAAAA 540
Db 725 ACCATGCCCTATGTTAAAGAAAGGAGATTTGGCCCTTGGATGGATAGCAGATAGAAAA 784
QY 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAGAGGATTTTCTCTCAGGA 600
Db 785 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAGAGGATTTTCTCTCAGGA 844
QY 601 TCTTTTGTCTGTATATTTCTGGCTTAAAGAAAGAGAGGCTTATGCGAGGAGCTCACTTTTTC 660
Db 845 TCTTTTGTCTGTATATTTCTGGCTTAAAGAAAGAGAGGCTTATGCGAGGAGCTCACTTTTTC 904
QY 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 905 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 964
QY 721 TACTTAGTAAATGAAGCTTTTAAAGAAAGGCTCAGAGAGATCATTTGATGCTGCAAA 780
Db 965 TACTTAGTAAATGAAGCTTTTAAAGAAAGGCTCAGAGAGATCATTTGATGCTGCAAA 1024
QY 781 ATTGACGAGGATTTTAAAGAAAGGCTTGCAGAGTGGCTCATTTGGAATGAATTGCATT 840
Db 1025 ATTGACGAGGATTTTAAAGAAAGGCTTGCAGAGTGGCTCATTTGGAATGAATTGCATT 1084

QY 841 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGGAATCTGGATCTTCA 900
DB 1085 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGGAATCTGGATCTTCA 1144
QY 901 AAGGTTTTTCAGCGCAAAATCCTTTTGTATTTTATGGAACATTTCTTTAGAGGAAAA 960
DB 1145 AAGGTTTTTCAGCGCAAAATCCTTTTGTATTTTATGGAACATTTCTTTAGAGGAAAA 1204
QY 961 ACAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAC 1020
DB 1205 ACAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAC 1264
QY 1021 ACAGATAACGTTCTTCACTTGGATGAGATTTT 1053
DB 1265 ACAGATAACGTTCTTCACTTGGATGAGATTTT 1297

RESULT 7

US-10-220-335-514/c
; Sequence 514, Application US/10220335
; Publication No. US20050175607A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-041
; CURRENT APPLICATION NUMBER: US/10/220,335
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 09/664,641
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/616,807
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 09/597,707
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/577,409
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: 09/515,126
; PRIOR FILING DATE: 2000-02-28
; NUMBER OF SEQ ID NOS: 688
; SOFTWARE: Custom
; SEQ ID NO 514
; LENGTH: 2482
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-220-335-514

Query Match 55.1%; Score 580; DB 24; Length 2482;
Best Local Similarity 75.1%; Pred. No. 4.5e-157;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;
QY 89 ATGAAGAGCCACTCTAAGAAAGAGTCTCGCGGTTTGTCTCATCTTCCAAATCCAGTACC 148
DB 2086 AGGATGAGCGGTCTGAGAGAAACCCCGCGCTTTGTCTCATCTTCCCATCGAGTACC 2027
QY 149 CTGATATTTGGAATAATGTATAACAGCGCAGGCTTCTTCTGACAGCAGAGAGGTTG 208
DB 2026 ATGATATCTGGCAGATGTATAGAAAGCAGAGCTTCTTTGGACCGCGAGAGGTTG 1967
QY 209 ACTTATCAAGAGATCTCCCTCACTGGAACAAAGCTTAAAGCAGATGAGAACTTCTCATCT 268
DB 1966 ACCTCTCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTATAT 1907
QY 269 CTCACATCTAGGCTTTTGGACCGAGTGTGGAATTTGAAATGAAATTTGGTGGAGC 328
DB 1906 CCCATGTTCTGGCTTTCTTTGCAAGCAGGATGGCATAGTAATGAAATCTTGGTGGAGC 1847
QY 329 GCTTTAGTCAGAGGTGCGAGTTCCAGAGCTCGCTGTTTCTATGCTTTCAAATCTCA 388
DB 1846 GATTTAGCCAAAGTTTCAGATTACAGAACCCGCTGTTTCTAAGCTTCCAAATGCGCA 1787
QY 389 TCGAATATCTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAGATCCCA 448
DB 1786 TGGAAACATACATCTGAAATGTATAGTCTTCTTATTGACACTTACATAAAGATCCCA 1727

QY 449 AGAAAGAGGAATTTTATTTAATGCAATTAAGAAACCATGCGCTATGTAAAGAAAAAGCAG 508
DB 1726 AGAAAGAGGAATTTTCTTCAATGCAATTTGAACGATGCTTGTGTCAAGAGAGGAGCAG 1667
QY 509 ATTGGCCCTTGGCATGGATAGCAGATAGAAAAATCTACTTTTGGGGAAGAGTGGTGGCCT 568
DB 1666 ACTGGCCCTTGGCATGGATTTGGGACAAAGAGGCTACCTATGTTGAAACGTTGTGTAGCCT 1607
QY 569 TTGCTGCTCTAGAGAGGATTTCTTCTCAGGATCTTTTGTCTGTATATCTTGGCTAAAGA 628
DB 1606 TTGCTGCTGAGGAGGATTTCTTCTTTCGGGTTCTTTTGGCTGATATTTCTGGCTCAAGA 1547
QY 629 AGAGAGGTCTTATGCGCAGGACTCACTTTTCCAAATGAACCTCATCAGCAGAGATCAAGGAC 688
DB 1546 AAGCAGGACTGATGCTTGGCTTCACTTTTCTAATGAACTTATTAGCAGAGATGAGGTT 1487
QY 689 TTTCACTGTGACTTTTGGCTTGGCTGATGTTTCAATTAATCTTGTAAATAAGCCCTTCAGAAAGAA 748
DB 1486 TACACTGTGATTTTGGCTTGGCTGATGTTTCAACACACCTGGTACACAAACCATCGGAGGAGA 1427
QY 749 GGGTCAGGAGATCATTGTTGATGCTGTCAAAATTTGAGCAGGAGTTTTAAACAGAACCT 808
DB 1426 GAGTAAGAGAAATAATTAATCAATGCTTTCGGATAGAACAGGAGTTTCTTCACTGAGGCT 1367
QY 809 TGGCAGTTGGCCCTCAATTTGGAATGAATTTGCAATTTTGTATGAACAGTACATTTGAGTTGTAG 868
DB 1366 TGGCTGTGAGCTCATTTGGATGATGTAATGCACTCTAATGAACTATACATTTGAGTTGTGG 1307
QY 869 CTGACAGATTTACTTTGGAACTTTGGAATTTCTCAAGGTTTTCAGGCGAGAAATCTTTTGG 928
DB 1306 CAGACAGACTTATGCTTGGAACTGGGTTTGTAGCAAGGTTTTCAGAGTAGAGAACCCATTTG 1247
QY 929 ATTTTATGAAACATTTCTTTAGAGGAAAAACAAATTTCTTTTGAACAGGAGTTTCAG 988
DB 1246 ACTTTATGAGATATTTTCACTGGAAGGAAAGACTAACTTTCTTTTGAAGAGAGTAGGCG 1187
QY 989 AGTATCAGCGCTTTTGCAGTTTATGCGAGAAACCAAGATTAAGCTTTCACCTTGGATGAG 1048
DB 1186 AGTATCAGAGGATGGAGTGTATCAAGTCAACAGAGAAATTTCTTTTACCTTGGATGCTG 1127
QY 1049 ATTT 1052
DB 1126 ACTT 1123

RESULT 8

US-09-954-456-724
; Sequence 724, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Can
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840

;
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 724
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-724

Query Match 55.1%; Score 580; DB 9; Length 2500;
Best Local Similarity 75.1%; Pred. No. 4.6e-157;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

QY 89 ATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGCGGGTTTGTCAATCTTCCAAATCCAGTACC 148
Db |||||
QY 397 AGGATGAGCGCTGTCTGAGAGAAACCCCGCGCTTTGTCAATCTTCCCATCGAGTACC 456
Db |||||
QY 149 CTGATATTTGGAAATGTATAACAGGCACAGCTTCTCTCGACAGCAGAGAGGTTG 208
Db |||||
QY 457 ATGATATCTGGCAGATGTATAAGAGCAGAGCTTCTTTTGGACCGCGAGAGGTTG 516
Db |||||
QY 209 ACTTATCAAGAGATCTCCCTCACTGGAAACAAGCTTAAAGCAGATGAGAAGTACTTTCATCT 268
Db |||||
QY 517 ACCTCTCCAAGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT 576
Db |||||
QY 269 CTCACATCTTAGCCTTTTTCAGCAGCAGTGTATGGAATGTAATGAAAAATTTGGTGGAGC 328
Db |||||
QY 577 CCCATGTTCTGGCTTTCTTTCAGCAAGCGATGGCATAGTAATGAAAACTTTGGTGGAGC 636
Db |||||
QY 329 GCTTTAGTCAGAGGTGCGGTTCCAGAGGCTCGCTGTTTCTATGCTTTCAAATTTCTCA 388
Db |||||
QY 637 GATTTAGCCAAAGTTTCAGATTCAGAAAGCCCGCTGTTTCTATGCGTTTCCAAATTTGCCA 696
Db |||||
QY 389 TCGAATATCTTCACTCAGAGATGTACAGTTTCTGTATGACACTTACATCAGAGATCCCA 448
Db |||||
QY 697 TCGAAACATACATCTGAAATGTATAGTCTTCTTATGACACTTACATAAAGATGCCA 756
Db |||||
QY 449 AGAAAGGGAAATTTTATTTAATGCAATGAAACCATGCCCTATGTTAAGAAAAAGCAG 508
Db |||||
QY 757 AAGAAGGGAAATTTCTTCAATGCCATTGAAACGATGCTTGTGTCAAGAAAGAGGCAG 816
Db |||||
QY 509 ATTGGCCCTTGGATGGATAGCAGATGAAATCTACTTTTGGGGAAAGAGTGGGCT 568
Db |||||
QY 817 ACTGGCCCTTGGCTGGATGGGGAACAAAGAGGCTACCTATGTTGAAACGCTTGTAGCT 876
Db |||||
QY 569 TTGCTGCTGTAGAAGGATTTTCTTCTCAGGATCTTTTCTGCTATATTTCTGCTAAAGA 628
Db |||||
QY 877 TTGCTGCTGTAGAAGGATTTTCTTCTTCCGGTTCTTTTCCGCGATATTTCTGCTCAAGA 936
Db |||||
QY 629 AGAGAGGCTTTATGCGAGGACTCAGCTTTTCCAAATGAACTCATCAGCAGAGATGAAGGAC 688
Db |||||
QY 937 AACGAGGACTGATGCTGCGCTCAGATTTTCTAATGAACTTATAGCAGAGATGAGGTT 996
Db |||||
QY 689 TTCACATGCACTTTCCTGCTGATGTTCCAAATCTTAACTAGTAATAGCCTTCAGAGAAA 748
Db |||||
QY 997 TACACTGTGATTTTGGCTTGGCTGATGTTCAACACACCTGTGACACAAACCATCGGAGGA 1056
Db |||||
QY 749 GGGTCAGGAGATCATTTGATGCTGTCAAAATTCAGCAGGATTTTAAACAGAGCCT 808
Db |||||
QY 1057 GAGTAAAGAAATTAATTAATCAATGCTGTTTGGATAGAACAGAGTTCTTCACTGAGGCT 1116
Db |||||
QY 809 TGCCAGTTGGCTTCATTGGAATGAATTCATTTTGAATGAAACAGTACATTTGATTTGTAG 868
Db |||||
QY 1117 TGCCTGTGAAGCTCATGGGATGATTCAGCTCTAATGAAGCAATACATTTGATTTGTGG 1176
Db |||||
QY 869 CTGACAGATTAATCTGTGAACTTGTGATTTCTCAAGGTTTTTTCAGGCAAGAAATTCCTTTG 928
Db |||||
QY 1177 CAGACAGACTTATGCTGAACTGGGTTTTAGCAAGGTTTTTCAGAGTAGAGAAACCCATTTG 1236
Db |||||
QY 929 ATTTTATCGAAACATTTCTTTAGAGGAAACAAATTTCTTTTGAAGAACAGATTTCAG 988
Db |||||

Db 1237 ACTTTATGGAGAATATTTTCACTGGAAGAAAGACTAACTTCTTTGAGAAGAGTAGGCG 1296
QY 989 AGTATCAGCGTTTTCAGTTATGCGCAAAACCAAGATAAAGTCTTTCACCTTGGATGAG 1048
Db |||||
QY 1297 AGTATCAGAGGATGGGAGTGTCTAAGTCCACAGAGAAATTTCTTTTACCTTGGATGCTG 1356
Db |||||
QY 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 9

US-09-954-456-1169
; Sequence 1169, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Can
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1169
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-1169

Query Match 55.1%; Score 580; DB 9; Length 2500;
Best Local Similarity 75.1%; Pred. No. 4.6e-157;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

QY 89 ATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGCGGGTTTGTCAATCTTCCAAATCCAGTACC 148
Db |||||
QY 397 AGGATGAGCGCTGTCTGAGAGAAACCCCGCGCTTTGTCAATCTTCCCATCGAGTACC 456
Db |||||
QY 149 CTGATATTTGGAAATGTATAAAGCAGCAGAGCTTCTCTCGACAGCAGAGAGGTTG 208
Db |||||
QY 457 ATGATATCTGGCAGATGTATAAAGCAGCAGAGCTTCTTTTGGACCGCGAGAGGTTG 516
Db |||||
QY 209 ACTTATCAAGAGATCTCCCTCACTGGAAACAAGCTTAAAGCAGATGAGAAGTACTTTCATCT 268
Db |||||
QY 517 ACCTCTCCAAGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT 576
Db |||||
QY 269 CTCACATCTTAGCCTTTTTCAGCAGCAGTGTATGGAATGTAATGAAAAATTTGGTGGAGC 328
Db |||||
QY 577 CCCATGTTCTGGCTTTCTTTCAGCAAGCGATGGCATAGTAATGAAAACTTTGGTGGAGC 636
Db |||||
QY 329 GCTTTAGTCAGAGGTGCGGTTCCAGAGGCTCGCTGTTTCTATGCTTTCAAATTTCTCA 388
Db |||||
QY 637 GATTTAGCCAAAGTTTCAGATTCAGAAAGCCCGCTGTTTCTATGCGTTTCCAAATTTGCCA 696
Db |||||

```
QY 389 TCGAAGATGTTCACTCAGAGATGTACAGTTTCTGCTATAGACACTTACATCAGAGATCCCA 448
Db 697 TGGAAACATACATCTCGAATGTAAGTCTTCTTATGACACTTACATAAAGATCCCA 756
QY 449 AGAAAGGGAATTTTATTAATGCAATGGAACCATGCCCTATGTTAAGAAAAAGCAG 508
Db 757 AAGAAAGGGAATTTCTCTCAATGCCATTGAAACGATGCTTGTCTCAAGAAGAGGCAG 816
QY 509 ATTGGGCTTCGGATGGATAGACATAGAAATCTACTTTTGGGGAAGAGATGGTGGCT 568
Db 817 ACTGGGCTTCGGATGGATGGGACAAAGAGGCTACCTATGCTGAACGTTGTGTAGCCT 876
QY 569 TTGCTGCTGTAGAGGATTTCTCTCAGGATCTTTTCTGCTATATTTCTGGCTAAAGA 628
Db 877 TTGCTGCTGTAGAGGATTTCTCTTCCGGTCTTTTGGCTGATATTTCTGGCTCAAGA 936
QY 629 AGAGAGTCTTATGCGAGGACTCACTTTTCCAAATGAACTCATCAGCAGAGATGAAGAC 688
Db 937 AACGAGGACTGATGCTGCGCTCACATTTTCTAATGAACTTATTAGCAGAGATGAGGTT 996
QY 689 TTCACTGTGACTTTGCTTGCCTGATGTTCCAAATACCTTAGTAATTAAGCCTTCAGAGAA 748
Db 997 TACACTGTGATTTTGTCTTGCCTGATGTTCAAAACACCTGCTACACAAACCATCGAGGAGA 1056
QY 749 GGGTCAGGAGATCATTGTTGATGCTGTCAAAATTCAGCAGGAGATTTTAAACAGAGCCT 808
Db 1057 GAGTAAGAGAAATAATTATCAATGCTGTTCGGATAGAACAGGAGTTCTCTCAGTGGCCT 1116
QY 809 TGCCAGTTGGCCTCATTTGGAATGAAATGCAATCTTCAAGGATTTTCAAGAGAGCCT 808
Db 1117 TGCCAGTTGGCCTCATTTGGAATGAAATGCAATCTTCAAGGATTTTCAAGAGAGCCT 1116
QY 869 CTGACAGATTAATGTTGGAATCTGGAATCTCAAGGTTTTCAGGAGAAATTCCTTTTG 928
Db 1177 CAGACAGACTTATGCTGGAATGCTGTTTTCAGGATGAAACAGGATTCCTCTCAGGAGCCT 1236
QY 929 ATTTTATGGAACATTTCTTTAGAGGAACAAATTTCTTTGAGAAACAGTTTTCAG 988
Db 1237 ACTTTATGGAATATTTTCTGGAAGGAAAGAACTAATCTTTTGAAGAGAGATGAGCG 1296
QY 989 AGTATCAGCTTTTTCAGTTATGCGAGAAACACAGATAAGCTTTCACTTGGATGCG 1048
Db 1297 AGTATCAGAGATGAGGATGATGCAAGTCCACAGAGAAATTTCTTTTACCTTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 10
US-09-954-456-1827
; Sequence 1827, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; PRIORITY FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
```

```
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 1827
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-954-456-1827

Query Match 55.1%; Score 580; DB 9; Length 2500;
Best Local Similarity 75.1%; Pred. No. 4.6e-157;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

QY 89 ATGAAGAGCCACTCTTAAGAAAGAGTTCTGCGCGGTTTGTTCATCTTCCATCCAGTACC 148
Db 397 AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTTGTTCATCTTCCCATCGAGTACC 456
QY 149 CTGATATTTGGAAATGTATAACAGGCACAGGCTTCCCTCTGGACACAGAGAGGTTG 208
Db 457 ATGATATCTGGCAGATGATATAAGAGGACAGAGGCTTCTTTGGACCCCGAGGAGGTTG 516
QY 209 ACTTATCAAGAGATCTCCCTCACTGGAAACAAGCTTAAAGCAGATGAGAAGTACTTCATCT 268
Db 517 ACCTCTCCAAGGACATTCAGCACTCGGAAATCCCTGAAACCCGAGGAGAGATATTTTATAT 576
QY 269 CTCACATCTTAGGCTTTTTCAGCCAGTGTAGGAATTTAAATGAAAAATTTGGTGGAGC 328
Db 577 CCATGTTCTGCGCTTTCTTTGACGAAAGCATAGTAAATGAAAACTTGGTGGAGC 636
QY 329 GCTTTAGTCAGAGGTCAGGTTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAATTCCTCA 388
Db 637 GATTTAGCCAGAGTTTCAGATTTACAGAGCCGCTGTTTCTATGGCTTCAAATTCCTCA 696
QY 389 TCGAGAATGTTCACTCAGAGATGTACAGTTTGTGTATAGACACTTACATCAGAGATCCCA 448
Db 697 TGGAAAAACATACATTCTGAAATGTATAGTCTTCTTATGACACTTACATAAAGATCCCA 756
QY 449 AGAAAGGGAATTTTATTTAATGCAATTTGAAACCATGCCCTATGTTAAGAAAAAGCAG 508
Db 757 AAGAAAGGGAATTTCTCTTCAATGCCATTTGAAACGATGCTTGTGTCAAGAAAGGAGCAG 816
QY 509 ATTGGGCTTCGGATGGATAGAGATAGAAATCTACTTTTGGGGAAGAGTGGTGGCCT 568
Db 817 ACTGGGCTTCGGCTGGATTTGGGACAAAGAGGCTACCTATGTTGAACGTTGTGTAGCCT 876
QY 569 TTGCTGCTGTAGAAGGAGTTTCTCTCAGGATCTTTTGTGCTGTATATTTCTGGCTAAAGA 628
Db 877 TTGCTGCTGTAGAAGGAGTTTCTTCTTCCGGTCTTTTGTGCTGATATTTCTGGCTCAAGA 936
QY 629 AGAGAGTCTTATGCGAGGACTCACTTTTCCAAATGAACTCATCAGCAGAGATGAAGAC 688
Db 937 AACGAGGACTGATGCTGCGCTCACATTTTCTAATGAACTTATTAGCAGAGATGAGGTT 996
QY 689 TTCACTGTGACTTTGCTTGCCTGATGTTCCAAATACCTTAGTAATTAAGCCTTCAGAGAA 748
Db 997 TACACTGTGATTTTGTCTTGCCTGATGTTCAAAACACCTGCTACACAAACCATCGAGGAGA 1056
QY 749 GGGTCAGGAGATCATTGTTGATGCTGTCAAAATTCAGCAGGAGATTTTAAACAGAGCCT 808
Db 1057 GAGTAAGAGAAATAATTATCAATGCTGTTCGGATAGAACAGGAGTTCTCTCAGTGGCCT 1116
QY 809 TGCCAGTTGGCCTCATTTGGAATGAAATGCAATCTTCAAGGATTTTCAAGAGAGCCT 808
Db 1117 TGCCAGTTGGCCTCATTTGGAATGAAATGCAATCTTCAAGGATTTTCAAGAGAGCCT 1176
QY 869 CTGACAGATTAATGTTGGAATCTGGAATCTCAAGGTTTTCAGGAGAAATTCCTTTTG 928
Db 1177 CAGACAGACTTATGCTGGAATGCTGTTTTCAGGATGAAACAGGATTCCTCTCAGGAGCCT 1236
QY 929 ATTTTATGGAACATTTCTTTAGAGGAACAAATTTCTTTGAGAAACAGTTTTCAG 988
Db 1237 ACTTTATGGAATATTTTCTGGAAGGAAAGAACTAATCTTTTGAAGAGAGATGAGCG 1296
QY 989 AGTATCAGCTTTTTCAGTTATGCGAGAAACACAGATAAGCTTTCACTTGGATGCG 1048
Db 1297 AGTATCAGAGATGAGGATGATGCAAGTCCACAGAGAAATTTCTTTTACCTTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 10
US-09-954-456-1827
; Sequence 1827, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; PRIORITY FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
```

Db 1177 CAGACAGACTTATGCTGGAACTGGGTTTATAGCAAGGTTTTCAGAGTAGAAGAACCCATTG 1236
Qy 929 ATTTTATGAAAAACATTTCTTTAGAAAGGAAAAACAAATTTCTTTTGAAGAACGAGTTTCAG 988
Db 1237 ACTTTATGAGAAATATTTCTCTGGAAGGAAGACTAACTTTCTTTGAGAGAGAGTAGGGC 1296
Qy 989 AGTATCAGCGTTTTCAGTTATGGCAGAAACACAGATAACGCTTTCACCTTCGATGCAG 1048
Db 1297 AGTATCAGAGGATGGGAGTGAAGTCAAGTCCAACAGAGAAATCTTTTACCTTCGATGCTG 1356
Qy 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 11

US-10-641-643-1370
; Sequence 1370, Application US/10641643
; Publication No. US20040077003A1
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; Susan G. Stuart
; Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
; GENE EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/641,643

FILING DATE: 14-Aug-2003

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: <Unknown>

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Zeller, Karen J.

REGISTRATION NUMBER: 37,071

REFERENCE/DOCKET NUMBER: PA-0001 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 855-0555

TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 1370:

SEQUENCE CHARACTERISTICS:

LENGTH: 2500 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GENBANK

CLONE: g36154

SEQUENCE DESCRIPTION: SEQ ID NO: 1370 :

US-10-641-643-1370

Query Match 55.1%; Score 580; DB 19; Length 2500;

Best Local Similarity 75.1%; Pred. No. 4.6e-157;

Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

Qy 89 ATGAAGACCCACTCTCTAAGAAAGAGTTCTCGCCGGTTTGTCTATCTTTTCCAATCCAGTACC 148

Db 397 AGGATGACCGCTGCTGAGAGAAACCCCGCGCTTTGTCTATCTTTCCCATCGAGTACC 456

Qy 149 CTGATATTTGGAAAAATGTATAAACAGGCACAGAGCTTCTCTGACAGCAGAGAGGTTG 208
Db 457 ATGATATCTGCAGATGTATAAGAAAGCAGAGGCTTCTCTTTGGACCCCGAGGAGTTG 516
Qy 209 ACTTATCAAAAGGATCTCCCTCAGCTGGAACAAAGCTTAAAGCAGATGAGAGTACTTCATCT 268
Db 517 ACCTCTCAAAGGACATTCAGCACTTGGGAATCCCTGAAACCCGAGGAGAGATATTTATAT 576
Qy 269 CTCACATCTTAGCCTTTTTCGACCCAGTGAATGTAATGTAATGTAATGTAATGTAATGTAAT 328
Db 577 CCCATGTTCTGGCTTCTTTTCGACCAAGCGATGTCATAGTAAATGAAACACTTGTGGAGC 636
Qy 329 GCTTTAGTCAGGAGGTGAGGTTCCAGAGGCTCGCTGTTCTATAGGCTTTCAAAATCTCA 388
Db 637 GATTTAGCCCAAGATTCAGATTACAGAGCCGCTGTTCTATGCGCTTCCAAATGCCA 696
Qy 389 TCGAGAAATGTTCACTCAGAGATGTACAGTTTGTGTATAGACACTTACATCAGAGATCCCA 448
Db 697 TGGAAAAACATACATTTCTGAAATGTATAGTCTTCTTATTGACACTTACATAAAGATCCCA 756
Qy 449 AGAAAAGGGAATTTTATTTAATGCAATTTGAACCATGCCCTATGTTAAGAAAAAGCAG 508
Db 757 AAGAAAAGGGAATTTCTCTTCAATGCCATTTGAACGATGCCCTGTTGTCAAGAAAGAGGAG 816
Qy 509 ATTGGGCTTGGCAGTAGCAGATAGAAAACTCTCTTTTGGGGAAGAGTGTGGGCT 568
Db 817 ACTGGGCTTGGCAGTAGGATTTGGGACAAAGAGGTACTATGTTGAACGTTGTAGGCT 876
Qy 569 TTGCTGCTGTAAGAGGAGTTTCTTCTCAGGATCTTTTGTGCTATATTTCTGGCTAAAGA 628
Db 877 TTGCTGAGTGAAGGAGATTTTCTTTTCCGGTCTTTTTCGCTGATATTTCTGGCTCAAGA 936
Qy 629 AGAGAGTCTTATGCCAGGACTCACTTTTCCAAATGAACCTCATCAGCAGAGATGAAGAC 688
Db 937 AACGAGGACTGATGCCCTGCTCACAATTTCTAATGAACCTTATTAGCAGAGATGAGGTT 996
Qy 689 TTCACGTGACTTTGCTTGGCTGATGTTCCAAATGTTAGTAAATGAAGCTTTCAGAGAA 748
Db 997 TACACTGTGATTTTGTCTGCTGATGTTCAAAACACCTGGTACACAAACCATCGGAGGAGA 1056
Qy 749 GGCTCAGGAGATCATTTGTTGATGCTGTCAAAATTTGAGCAGGAGTCTTTTAAACAGAGCT 808
Db 1057 GAGTAAGAGAAATTAATCAATGCTGTTCCGATAGAACAGGAGTCTCTCACTGAGGCT 1116
Qy 809 TGCCAGTTGGCCTCATTTGGAATGAATGCAATTTGATGAAACAGTACATTGAGTTGTAG 868
Db 1117 TGCTGTGAAGCTCATTTGGGATGAATTCGACTCTAATGAAGCAATACATTGAGTTGTGG 1176
Qy 869 CTGACAGATTACTTTGGGACTTGGATTTCAAAAGGTTTTTCAGGCGAGAAATCTTTTG 928
Db 1177 CAGACAGACTTATGCTGGAACTGGGTTTATAGCAAGGTTTTTTCAGAGTAGAGAACCATTTG 1236
Qy 929 ATTTTATGAAAAACATTTCTTTTGAAGGAAAAACAAATTTCTTTTGAAGAACGAGTTTCAG 988
Db 1237 ACTTTATGAGATATTTTCTCGAAGGAAGAGTAACTTCTTTTGAAGAGAGTAGCGG 1296
Qy 989 AGTATCAGCGTTTTCAGTTATGGCAGAAACCAAGATAACGCTTTCACCTTCGATGCAG 1048
Db 1297 AGTATCAGAGGATGGGAGTGTCAAGTCCAACAGAGAAATCTTTTACCTTGGATGCTG 1356
Qy 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 12

US-10-733-878-458
; Sequence 458, Application US/10733878
; Publication No. US20040224408A1
; GENERAL INFORMATION:
; APPLICANT: Jean-Philippe Girard
; APPLICANT: Francois Amalric
; APPLICANT: Myriam Roussigne

APPLICANT: Thomas Clouaire
TITLE OF INVENTION: THAP PROTEINS AS NUCLEAR RECEPTORS FOR
TITLE OF INVENTION: CHEMOKINES AND ROLES IN TRANSCRIPTIONAL REGULATION, CELL
TITLE OF INVENTION: PROLIFERATION AND CELL DIFFERENTIATION
FILE REFERENCE: BIOBANK.012A
CURRENT APPLICATION NUMBER: US/10/733.878
CURRENT FILING DATE: 2003-12-10
PRIOR APPLICATION NUMBER: 60/432699
PRIOR FILING DATE: 2002-12-10
PRIOR APPLICATION NUMBER: 60/485027
PRIOR FILING DATE: 2003-07-03
NUMBER OF SEQ ID NOS: 535
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 458
LENGTH: 2500
TYPE: DNA
ORGANISM: Homo sapiens
US-10-733-878-458

Query Match 55.1%; Score 580; DB 21; Length 2500;
Best Local Similarity 75.1%; Pred. No. 4.6e-157;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

89 ATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGCGGTTTGTTCATCTTTTCCAATCCAGTACC 148
397 AGGATGAGCCGCTGCTGAGAGAAACCCCGCGCTTTGTCTATCTTCCCATCGAGTACC 456
149 CTGATATTGGAAATATGATATAACAGGCACAGGCTTCTCTGACAGCAGAAAGAGTTG 208
457 ATGATATCTGCAGATGATATAAGAGGCGAGGCTTCTTTTGGACCGCGAGGAGTTG 516
209 ACTTATCAAGAGTCTCCCTCACTGGAACAGCTTAAAGCAGATGAGAGTACTTCATCT 268
517 ACCTCTCAAGAGCATTCAGCACTGGGAATCCCTGAAACCCGAGGAGATATTTTATAT 576
269 CTCACATCTTAGCCCTTTTTCAGCCAGTGTGAATTTGAAATTTGTTGGAGC 328
577 CCCATGTTCTGCTTCTTTGACGACGAGGATGATAGTAATGAAATTTGTTGGAGC 636
329 GCTTTAGTTCAGAGGTGCGTTTCAGAGGCTGCTGTTTCTATGGCTTTCAAAATCTCA 388
637 GATTTAGCAAGAAGTTCCAGATTAAGAAAGCCGCTGTTTCTATGGCTTCAAAATTCGA 696
389 TCGAGATGTTCACTCAGAGATGATACGTTTCTGCTATGACACTTACATCAGAGTCCCA 448
697 TGGAAACATACATCTGAAATGATAGTCTTCTTATGACACTTACATNAAGATCCCA 756
449 AGAAAGGGAATTTTATTTAATGCAATTTGAAACCATGCTTATGTTAAGAAAGAGCAG 508
757 AAGAAAGGGAATTTCTCTCAATGCCATTGAAACGATGCTTGTCAAGAAAGAGGAG 816
509 ATTGGGCTTCGATGAGATGAGATGAAATCTACTTTTGGGAAAGAGTGTGGCT 568
817 ACTGGGCTTCGCTGGATTTGGGACAAAGAGGCTACCTATGTTGAACGTTGTAGCCT 876
569 TTGCTGCTGATAGAGGAGTTTCTTCTCAGGATCTTTTCTGCTATATCTTGGCTAAAGA 628
877 TTGCTGCTGATGAGAGGCAATTTCTTTTCCGGTTCTTTTCCGTCGATATCTTGGCTCAAGA 936
629 AGAGAGGCTTTATGCGAGGACTCACTTTTCCAATGAACTCATCAGCAGAGATGAAGAC 688
937 AACGAGGACTGATGCTGCTGCTCACTTTTCTAATGAACTTATAGCAGAGATGAGGTT 996
689 TTCACTGTGACTTTGCTTTCCTGATGTTTCCAATCTTATAGTAATAGGCTTTCAAGAGAAA 748
997 TACACTGTGATTTTGTCTTGGCTGATGTTTCAACACCTGCTACACAAACCATCGGAGGA 1056
749 GGGTCAGGAGATCATTTGATGCTGTCAAAATTTAGCAGAGGTTTAAACAGAGGCT 808
1057 GAGTAAGAGAAATAATATCAATGCTGTTTCGATAGAAACAGAGGTTCTCTCACTGAGGCT 1116
809 TGCCAGTTGGCTCATTTGGAATGATTCATTTTGAATGAAACAGATGAGATTCAGTTGTAG 868

Db 1117 TGCCTGTGAAGCTCATTTGGGATGAATTCACCTCTATGAAGCAATACATTGAGTTTGTGG 1176
Qy 869 CTGACAGATTACTTGTGGAACTTGAATCTCAAGGTTTTCAGGCAAGAAATCCTTTTG 928
Db 1177 CAGACAGACTTATGCTGGAACTGGGTTTATAGCAAGGTTTTCAGAGTAGAAGCCCATTTG 1236
Qy 929 ATTTTATGAAAAACATTTCTTTAGAGGAAAAACAAATTTCTTTGAGAAACGAGTTTCAG 988
Db 1237 ACTTTATGGAGATATTTTCACTGGAGGAAAGACTTAATCTTTTGAAGAGAGTAGCG 1296
Qy 989 AGTATCAGCGTTTTCAGTTATGCGAGAAACACAGATAACGCTTTCACCTTGGATGACAG 1048
Db 1297 AGTATCAGAGGATGGAGTGTCAAGTCCAACAGAGAAATCTTTTACCTTGGATGCTG 1356
Qy 1049 ATTTT 1052
Db 1357 ACTT 1360

RESULT 13

US-10-843-641A-3751
Sequence 3751, Application US/10843641A
Publication No. US20050064454A1
GENERAL INFORMATION:
APPLICANT: Avalon Pharmaceuticals, Inc.
TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
TITLE OF INVENTION: Signature Gene Sets
FILE REFERENCE: 689290-189
CURRENT APPLICATION NUMBER: US/10/843,641A
CURRENT FILING DATE: 2004-05-12
PRIOR APPLICATION NUMBER: US/09/873,367
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: US/09/954,531
PRIOR FILING DATE: 2001-09-18
PRIOR APPLICATION NUMBER: US/09/954,456
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US/09/962,436
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US/09/962,832
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US/09/964,824
PRIOR FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: US/09/967,768
PRIOR FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: US/09/968,007
PRIOR FILING DATE: 2001-10-02
PRIOR APPLICATION NUMBER: US/09/969,347
PRIOR FILING DATE: 2001-10-02
PRIOR APPLICATION NUMBER: US/09/969,708
PRIOR FILING DATE: 2001-10-03
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 847
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3751
LENGTH: 2500
TYPE: DNA
ORGANISM: Homo sapiens
US-10-843-641A-3751

Query Match 55.1%; Score 580; DB 22; Length 2500;
Best Local Similarity 75.1%; Pred. No. 4.6e-157;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

89 ATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGCGGTTTGTTCATCTTTTCCAATCCAGTACC 148
397 AGGATGAGCCGCTGCTGAGAGAAACCCCGCGCTTTGTTCATCTTCCCATCGAGTACC 456
149 CTGATATTGGAAATATGATATAACAGGCACAGGCTTCTTCTGACAGCAGAAAGAGTTG 208
457 ATGATATCTGCGAGATGATATAAGAGGCGAGGCTTCTTTTGGACCGCGAGGAGTTG 516
209 ACTTATCAAGAGTCTCCCTCACTGGAACAGGTTTAAAGCAGATGAGAGTACTTCATCT 268

```

Db 517 ACCTCTCCAAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT 576
QY 269 CTCACATCTTAGCCTTTTTCAGCCAGGATGGAATCTAATGAATTTGGTGGAGC 328
Db 577 CCATGTTCTGCTTCTTTGAGCAAGGATGGAATGAATGAATTTGGTGGAGC 636
QY 329 GCTTTAGTCAGGAGTGCAGGTTCCAGAGGCTCGCTGTTTCTATGCTTTTCAAAATCTCA 388
Db 637 GATTTAGCCAAAGTTCAGATTACAGAAGCCGCTGTTCTATGGCTTCCAAATGGCA 696
QY 389 TCGAATGTTCTCAGAGATGTACAGTTGCTGATGACACTTACATCAGAGATCCCA 448
Db 697 TCGAATAACATACATTCGAATGTATAGTCTCTTATTGACACTTACATAAAGATCCCA 756
QY 449 AGAAAGGGAATTTTATTATTAATGCAATGAACCATGCCCTATGTTAGAAAGGAGC 508
Db 757 AGAAAGGGAATTTCTCTCAATGCCATTGAACCATGCCCTGTTGTCAAGAGAGAGGCG 816
QY 509 ATTGGGCTTGGATGGATAGCAGATAGAAATCTACTTTTGGGAAAGAGTGGTGGCT 568
Db 817 ACTGGGCTTGGCTGATGGATGGGCAAGAGGCTACCTATGTAACGTGTTGTAGCT 876
QY 569 TTGCTGCTAGAGGATTTCTCTCAGGATCTTTTCTCAGGATCTTTTCTGCTATATTTCTGCTAAAGA 628
Db 877 TTGCTGCTAGGAGGATTTCTTTTCCGGTCTTTTCCGGTCTTTTCCGGTCTGATATTTCTGCTCAAGA 936
QY 629 AGAGAGTCTTATGCCAGGACTCACTTTTCCAATGAACTATCAGCAGAGATGAAGGAC 688
Db 937 AACGAGGACTGATGCTGCGCTCCATTTTCTAATGAATTTATTAGCAGAGATGAGGTT 996
QY 689 TTCCTGTGACTTGTGCTTCCATGATGTTTCCATGATGTTTAAATAGGCTTTCAGAGAAA 748
Db 997 TACATGATTTTGTCTGCTGATGTTTCAACACCTGGTACACAAACCTCGGAGAGA 1056
QY 749 GGGTCAGGAGATCAATTTGATGCTGCTCAAAATTTGAGCAGGAGTTTTCAGAGAGCCT 808
Db 1057 GAGTAAGAGAAATATTAATCAATGCTGTTCCGATAGAACAGGTTTCTCTAGGCT 1116
QY 809 TGCCAGTTGGCTCATTTGAATGAATTTGATTTGATGAAACAGTACATTTGATGTTGATG 868
Db 1117 TGCTGTGAAAGCTCATTTGGATGAATTTGACCTCTAATGAAGCAATACATTTGATTTGTTG 1176
QY 869 CTGACAGATTACTTGTGAACTTGTGATTTCTCAAGGTTTTCAGCAGAGAAATCTTTTG 928
Db 1177 CAGACAGACTTATGCTGAACTGGGTTTATAGCAAGTTTTCAGAGTAGAGAAACCCATT 1236
QY 929 ATTTTATGGAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACGAGTTTTCAG 988
Db 1237 ACTTTATGAGNATATTTCTGGAAGGAAAGACTTAACTTTTGGAGAGAGATAGGCG 1296
QY 989 AGTATCAGGTTTTCAGTTATGCGAGAAACCAAGATTAACGCTTTCACCTTTGGATGCG 1048
Db 1297 AGTATCAGAGATGGGATGATGTCAGAGTCCAAACAGAGAAATTTCTTTTACCTTTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360

```

```

RESULT 14
US-10-843-641A-4196
; Sequence 4196, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals, Inc.
; TITLE OF INVENTION: Gene Determination and Therapeutic Screening Using
; FILE OF INVENTION: Signature Gene Sets
; FILE REFERENCE: 689290-189
; CURRENT APPLICATION NUMBER: US/10/843,641A
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/873,367
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US/09/954,531

```

```

; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,832
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/964,824
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/09/967,768
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/09/968,007
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,708
; PRIOR FILING DATE: 2001-10-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 8447
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4196
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-843-641A-4196

```

```

Query Match 55.1%; Score 580; DB 22; Length 2500;
Best Local Similarity 75.1%; Pred. No. 4.6e-157;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

```

```

QY 89 ATGAAGAGCCACTCTTAAGAAGAGTCTCGCGGTTTGTCTCTCTGAGCAGAGAGAGGTTG 148
Db 397 AGATGAGCGCTGCTGAGAGAGAAACCCCGCGCTTTGTCTCTCTCTCTCTCTCTCTCTCT 456
QY 149 CTGATATTTGGAAATGTATAAAGCAGCAGAGCTTCTCTCTGAGCAGAGAGAGAGGTTG 208
Db 457 ATGATATCTGGCAGATGTATAAAGCAGCAGAGCTTCTCTCTGAGCAGAGAGAGGTTG 516
QY 209 ACTTATCAAGAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 268
Db 517 ACCTCTCCAAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT 576
QY 269 CTCACATCTTAGCCTTTTTCAGCCAGGATGGAATTTAAATGAAGATTTGTTGAGC 328
Db 577 CCATGTTCTGCTTCTTTTTCAGCAAGGATGGAATTTAAATGAAGATTTGTTGAGC 636
QY 329 GCTTTAGTCAGGAGTGCAGGTTTTCAGAGGCTCGCTGTTTCTATGCTTTTCAAAATCTCA 388
Db 637 GATTTAGCCAAAGTTCAGATTACAGAAGCCGCTGTTTCTATGGCTTCCAAATGGCA 696
QY 389 TCGAATGTTCTCAGAGATGTACAGTTTGTGATGACACTTACATCAGAGATCCCA 448
Db 697 TCGAATAACATACATTCGAATGTATAGTCTCTTATTGACACTTACATAAAGATCCCA 756
QY 449 AGAAAGGGAATTTTATTATTAATGCAATTTGAACCATGCCCTATGTTAGAAAGGAGC 508
Db 757 AGAAAGGGAATTTCTCTCAATGCCATTGAACCATGCCCTGTTGTCAAGAGAGGCG 816
QY 509 ATTGGGCTTGGATGGATAGCAGATAGAAATCTACTTTTGGGAAAGAGTGGTGGCT 568
Db 817 ACTGGGCTTGGCTGATGGATGGGCAAGAGGCTACCTATGTAACGTGTTGTAGCT 876
QY 569 TTGCTGCTAGAGGATTTTCTCTCAGGATCTTTTCTGCTGATATTTCTGCTTAAAGA 628
Db 877 TTGCTGCTAGGAGGATTTTCTTTTCCGGTCTTTTCCGGTCTTTTCCGGTCTGATATTTCTGCTCAAGA 936
QY 629 AGAGAGTCTTATGCCAGGACTCACTTTTCCAATGAACTATCAGCAGAGATGAAGGAC 688
Db 937 AACGAGGACTGATGCTGCGCTCCATTTTCTAATGAATTTATTAGCAGAGATGAGGTT 996
QY 689 TTCCTGTGACTTGTGCTTCCATGATGTTTCCATGATGTTTAAATAGGCTTTCAGAGAAA 748
Db 997 TACATGATTTTGTCTGCTGATGTTTCAACACCTGGTACACAAACCTCGGAGAGA 1056

```


Result No.	Score	Query %		DB	ID	Description
		Match	Length			
1	1051.4	99.8	3292	3	CR617553	full-length
2	1051.4	99.8	4748	3	HSW802033	AL137348 Homo sapi
3	1039.4	98.7	4550	3	BC042948	BC042948 Homo sapi
4	1009	95.8	3501	3	CR627376	CR627376 Homo sapi
5	943.4	89.6	1056	9	AY398973	AY398973 Homo sapi
6	928	88.1	1083	1	AL547501	AL547501 AL547501
7	828.2	78.7	1008	9	AY398974	AY398974 Pan trogl
8	791.6	75.2	879	5	BQ441857	BQ441857 AGENCOURT
9	778.6	73.9	1043	9	AY398975	AY398975 Mus muscu
10	731.2	69.4	892	7	CL1633214	CL1633214 952695 MA
11	726.6	69.0	914	7	CF9595079	CF9595079 AGENCOURT
12	710.6	67.5	870	7	CO648380	CO648380 ILLUMIGEN
13	682.8	64.8	804	6	CD565661	CD565661 AGENCOURT
14	653.8	62.1	1061	4	BM801298	BM801298 AGENCOURT
15	653.2	62.0	896	5	BUI70979	BUI70979 AGENCOURT
16	614.8	58.4	685	5	BX951374	BX951374 DKFzp781E
17	578.4	54.9	1571	3	CR603461	CR603461 full-length
18	578.4	54.9	1573	3	CR625489	CR625489 full-length
19	578.4	54.9	1582	3	CR608076	CR608076 full-length
20	578.4	54.9	1588	3	CR602054	CR602054 full-length
21	578.4	54.9	1592	3	CR621427	CR621427 full-length
22	578.4	54.9	1600	3	CR604378	CR604378 full-length
23	578.4	54.9	1605	3	CR590959	CR590959 full-length
24	578.4	54.9	1612	3	CR614990	CR614990 full-length

QY	61	GACACCAACGAAGTGAATTAAGTCAATTAAGAGCCACTCTTAAGAAAGATTCTCGC	120
Db	87	GACACCAACGAAGTGAATTAAGTCAATTAAGAGCCACTCTTAAGAAAGATTCTCGC	146
QY	121	CGGTTTGTTCATCTTTCCCAATCAGTACCTCTGATATTGGAAATGTATAAACAGGCACAG	180
Db	147	CGGTTTGTTCATCTTTCCCAATCAGTACCTCTGATATTGGAAATGTATAAACAGGCACAG	206
QY	181	GCTTCCTTCTGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAAACAG	240
Db	207	GCTTCCTTCTGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAAACAG	266
QY	241	CTTAAGCAGATGAGAGTACTTCTCTCATCTTAGCTTTTTCAGCCAGTGTAT	300
Db	267	CTTAAGCAGATGAGAGTACTTCTCTCATCTTAGCTTTTTCAGCCAGTGTAT	326
QY	301	GGAATTGTAATGAAATTTGGTGAGCGCTTTAGTCAGAGGTGCGAGTTCAGAGGCT	360
Db	327	GGAATTGTAATGAAATTTGGTGAGCGCTTTAGTCAGAGGTGCGAGTTCAGAGGCT	386
QY	361	CGCTTTTCTATGGTTTCAATTTCTCATCGAATGTTTCACTCAGAGATGTACAGTTTG	420
Db	387	CGCTTTTCTATGGTTTCAATTTCTCATCGAATGTTTCACTCAGAGATGTACAGTTTG	446
QY	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAAATCAATTGAA	480
Db	447	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAAATCAATTGAA	506
QY	481	ACCATGCCCTATGTTAAGAAAAAGCAGATTGGCGCTTCGATGATAGATAGATAAAA	540
Db	507	ACCATGCCCTATGTTAAGAAAAAGCAGATTGGCGCTTCGATGATAGATAGATAAAA	566
QY	541	TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGTCTTCTCTCAGGA	600
Db	567	TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGTCTTCTCTCAGGA	626
QY	601	TCCTTTGCTGCTATATTTCTGGCTTAAGAAAGAGGCTTATGCCAGGACTCATTTTTC	660
Db	627	TCCTTTGCTGCTATATTTCTGGCTTAAGAAAGAGGCTTATGCCAGGACTCATTTTTC	686
QY	661	AATGAACATCATCAGCAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA	720
Db	687	AATGAACATCATCAGCAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA	746
QY	721	TACTTTAGTAAATTAAGCCTTTCAGAAAGAAAGGTCAGGAGATCAITTTGATGCTGTCAA	780
Db	747	TACTTTAGTAAATTAAGCCTTTCAGAAAGAAAGGTCAGGAGATCAITTTGATGCTGTCAA	806
QY	781	ATTGAGCAGGAGTTTAAACAGAAAGCTTGCAGTTGGCTCATTTGGAATGAATTGCATT	840
Db	807	ATTGAGCAGGAGTTTAAACAGAAAGCTTGCAGTTGGCTCATTTGGAATGAATTGCATT	866
QY	841	TTGATGAACACGTACATTGAGTTTGTAGCTGACAGATTACTTGTGAACTTGCATTCTCA	900
Db	867	TTGATGAACACGTACATTGAGTTTGTAGCTGACAGATTACTTGTGAACTTGCATTCTCA	926
QY	901	AAGGTTTTTTCAGGCGAATAATCCTTTTGTATTTTATGAAAAACATTTCTTTAGAGAAAA	960
Db	927	AAGGTTTTTTCAGGCGAATAATCCTTTTGTATTTTATGAAAAACATTTCTTTAGAGAAAA	986
QY	961	ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC	1020
Db	987	ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC	1046
QY	1021	ACAGATAAGCTTCTTACCTTGCATCGAGATTTT	1053
Db	1047	ACAGATAAGCTTCTTACCTTGCATCGAGATTTT	1079
RESULT 2			
HSM802033			
LOCUS			
HSM802033			
4748 bp			
mRNA			
linear			
HTC 22-SEP-2004			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
CONSRM			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
1. .4748			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="RZPD:DKFZp761E1312"			
/db_xref="taxon:9606"			
/clone="DKFZp761E1312"			
/tissue type="amygdala"			
/clone_lib="761 (synonym: hamy2). Vector pSport1; host			
DH10B; sites NotI + SalI"			
/dev_stage="adult"			
/note="ribonucleotide reductase"			
1. .4748			
/genes="DKFZp761E1312"			
47. .1102			
/gene="DKFZp761E1312"			
/codon_start=1			
/product="hypothetical protein"			
/protein_id="CAB70703.2"			
/db_xref="GI:52545612"			
/db_xref="GOA:Q9NTD8"			
/db_xref="UniProt/TrEMBL:Q9NTD8"			
/translation="MGDPPEAAGLDQDERSSDTNESIKNEEPLLRKSSRRFVI			
FPIQYDPIWMYKQAQASFTWAEVDLSKDLPHWNKLKADKPYFISHILAFPAASDGI			
VNNELVERFSQVPEARCFYGFQILINHVHSEMSLLIDTVIRDPKREFLFLNAIE			
TWPYVKKADWLRWIADRKSTFGRVWAFVAAVEGVFSGKFAAIPLKKRGLMPGLT			
FSNELSRDEGLHCDFACLMFOVLNKPSEERVREIIVDAVKIQEQLTALPVGLIG			
MNCILMKQIIEFVADRLVLELGFVSKVQAEINPFDMENISLEKTNFPEKRVSEYQRF			
AWAETTDNDVFTLDADF"			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 1052; Conservative			
99.8%; Score 1051.4; DB 3; Length 4748;			
99.9%; Pred. No. 1.6e-268;			
0; Mismatches 1; Indels 0; Gaps 0;			
QY	1	ATGGCGCACCCGAAAGCGCGGCGCGGCTGGATCAGGATGAGAGATCATCTTCA	60
Db	47	ATGGCGCACCCGAAAGCGCGGCGCGGCTGGATCAGGATGAGAGATCATCTTCA	106
QY	61	GACACCAACGAAGTGAATTAAGTCAATTAAGAGCCACTCTTAAGAAAGATTCTCGC	120
Db	107	GACACCAACGAAGTGAATTAAGTCAATTAAGAGCCACTCTCTTAAGAAAGATTCTCGC	166
QY	121	CGGTTTGTTCATCTTTCCCAATCAGTACCTCTGATATTGGAAATGTATAAACAGGCACAG	180
Db	167	CGGTTTGTTCATCTTTCCCAATCAGTACCTCTGATATTGGAAATGTATAAACAGGCACAG	226


```
QY 181 GCTTCCTTCTGACAGCAGAGAGGTGACTTATCAAGGATCTCCCTCACTGGACAACAG 240
Db |||||
QY 227 GCTTCCTTCTGACAGCAGAGAGGTGACTTATCAAGGATCTCCCTCACTGGACAACAG 286
Db |||||
QY 241 CTTAAAGCAGATGAGAAGTACTTCATCTCTCATCTTACATCTTTAGCCCTTTTTCAGCCAGTGAT 300
Db |||||
QY 287 CTTAAAGCAGATGAGAAGTACTTCATCTCTCATCTTACATCTTTAGCCCTTTTTCAGCCAGTGAT 346
QY 301 GGAATTGTAATGAAATTTGGTGAGCGCTTTAGTCAGGAGTGCAGGTTCAGAGGCT 360
Db |||||
QY 347 GGAATTGTAATGAAATTTGGTGAGCGCTTTAGTCAGGAGTGCAGGTTCAGAGGCT 406
QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAAGTGTTCATCTCAGAGATGTACAGTTTG 420
Db |||||
QY 407 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAAGTGTTCATCTCAGAGATGTACAGTTTG 466
QY 421 CTGATGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTGAA 480
Db |||||
QY 467 CTGATGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTGAA 526
QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGGCTTCGCGATGATGACAGATGAGAAA 540
Db |||||
QY 527 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGGCTTCGCGATGATGACAGATGAGAAA 586
QY 541 TCTACTTTTGGGAAAGAGTGTGCGCTTTGCTGTGTAGAGGAGTTTCTTCTCAGGA 600
Db |||||
QY 587 TCTACTTTTGGGAAAGAGTGTGCGCTTTGCTGTGTAGAGGAGTTTCTTCTCAGGA 646
QY 601 TCTTTGCTGTATATCTCGCTAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 660
Db |||||
QY 647 TCTTTGCTGTATATCTCGCTAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 706
QY 661 AATGAACCTATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTGTGATGTTTCCAA 720
Db |||||
QY 707 AATGAACCTATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTGTGATGTTTCCAA 766
QY 721 TACTTAGTAATAGCCTTCAGAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAA 780
Db |||||
QY 767 TACTTAGTAATAGCCTTCAGAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAA 826
QY 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCCAGTTGGCTCATTTGGAATGAATTCATT 840
Db |||||
QY 827 ATTGAGCAGGAGTTTAAACAGAGCCTTGCCAGTTGGCTCATTTGGAATGAATTCATT 886
QY 841 TTGATGAACAGTACATTGAGTTTGTAGTCAGCAGATTTCTGTGGAACCTTTGGAATTC 900
Db |||||
QY 887 TTGATGAACAGTACATTGAGTTTGTAGTCAGCAGATTTCTGTGGAACCTTTGGAATTC 946
QY 901 AAGTTTTTTCAGCAGAAAATCCCTTTGATTTTATGGAACATTTCTTTAGAGAGAAA 960
Db |||||
QY 947 AAGTTTTTTCAGCAGAAAATCCCTTTGATTTTATGGAACATTTCTTTAGAGAGAAA 1006
QY 961 ACAAATTTCTTTGAGAAACAGATTTTCAGATATCAGCGTTTTCAGGTTATGCGAGAAACC 1020
Db |||||
QY 1007 ACAAATTTCTTTGAGAAACAGATTTTCAGATATCAGCGTTTTCAGGTTATGCGAGAAACC 1066
QY 1021 ACAGATAACGCTTTTCACTTCGGATGACAGATTTT 1053
Db |||||
QY 1067 ACAGATAACGCTTTTCACTTCGGATGACAGATTTT 1099
```

```
RESULT 3
BC042948
LOCUS BC042948 Homo sapiens cdna clone IMAGE:4798175, containing frame-shift errors.
DEFINITION BC042948 Homo sapiens cdna clone IMAGE:4798175, containing frame-shift errors.
ACCESSION BC042948
VERSION BC042948.2 GI:34194000
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

REFERENCE
AUTHORS

Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.

1 (bases 1 to 4650)
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Wotley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A., Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield J.S., Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E., Jones S.J. and Marra M.A.

TITLE

JOURNAL
PUBMED

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932

2 (bases 1 to 4650)
Director MGC Project.
Direct Submission
Submitted (09-JAN-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REFERENCE
AUTHORSTITLE
JOURNAL

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
On Aug 25, 2003 this sequence version replaced gi:27695575.
Contact: MGC help desk
Email: cgabbs@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca

REMARK

COMMENT

Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prabh, Parvaneh Saeedi, JR Santos, Angelique Schnerch, Ursula Skalska, Duane Smailus, Jeff Stott, Miranda Tsai, George Yang, Jacques Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK L.M.A.G.E. Row: h Column: 19
This clone has the following problem: frame shifted.

FEATURES
source

1..4650
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4798175"
/tissue_type="Brain, hypothalamus"
/clone_lib="NIH MGC_96"
/lab_host="DH10B"
/note="Vector: pBluescript"

ORIGIN

Query Match 98.7%; Score 1039.4; DB 3; Length 4650;
Best Local Similarity 99.8%; Pred. No. 2.5e-265;
Matches 1051; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY

1 ATGGGGACCCGAAAGCGCGAGCGCGCTCGATCAGGATGAGATCATCTTCA 60
|||||

LOCUS	CR627376	3501 bp	mRNA	linear	HTC 22-SEP-2004
DEFINITION	Homo sapiens mRNA; cDNA DKFZp686M05248 (from clone DKFZp686M05248).				
ACCESSION	CR627376				
VERSION	CR627376.1	GI:50949847			
KEYWORDS	HTC.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
	1 (bases 1 to 3501)				
AUTHORS	Ottenwaelder,B., Obermaier,B., Deutschenbaur,S., Schaipp,A.,				
	Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and				
	Wiemann,S.				
CONSTRM	The German cDNA Consortium				
	Direct Submission				
	Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764				
JOURNAL	Neuherberg, GERMANY				
	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer				
	Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;				
COMMENT	sequenced by Medigenomix (Martinsried/Germany) within the cDNA				
	sequencing consortium of the German Genome Project.				
	This clone (DKFZp686M05248) is available at the RZPD Deutsches				
FEATURES	Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.				
	Please contact RZPD for ordering:				
	http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp686M05248				
source	Further information about the clone and the sequencing project is				
	available at http://mips.gsf.de/projects/cdna/.				
	Location/Qualifiers				
gene	1..3501				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
CDS	/db_xref="taxon:9606"				
	/clone="DKFZp686M05248"				
	/tissue_type="salivary gland"				
ORIGIN	/clone_lib="686 (synonym: hlcc3) . Vector pSport1_sfi; host				
	DH10B; sites SfiIA + SfiIB"				
	/dev_stage="adult"				
Query Match	/note="ribonucleotide reductase M2 B (TP53 inducible)"				
	1..3501				
	/gene="DKFZp686M05248"				
Best Local Similarity	332..1225				
	/gene="DKFZp686M05248"				
	/codon_start=1				
Matches 1018; Conservative 0; Mismatches 15; Indels 0; Gaps 0;	/product="hypothetical protein"				
	/protein_id="CAH10473.1"				
	/db_xref="GI:50949848"				
QY	/translation="MYKQAQSFWTAEVDLSKDLPHWNKLKADERYFISHILAFPA				
	SDGIYVENLVERFSEVOVPEARCFYGOILLENVHSEMYSLIIDTVIRDPKKRFLF				
	NAIETMPYVKKKADWALRWIADRKSTFGERVVAFVAVGVPFSGSFAAI FPKKRGFLM				
DB	PGLTFSNELISRDEGLHCDPACLMFOYLVNKPPEERVREIIVDAVKLEQBELPEALPV				
	GLIGNWCILMKQYIEFVADRLLLVGFSGVSKVQAEVNFDFMENISLEKTNPFPEKRVSE				
	YQRFVMAETTDNVFTLDADF"				
QY	21 GGAAGCGCGCGCTGGATCAGGATGAGAGATCATCTTCAGACACCAACGAAAGTGAAT				
	80				
	190 GGACGACGCTTGAGGCGCTCAGGCCGGGAGATCATCTTCAGACACCAACGAAAGTGAAT				
QY	81 AAGTCAAAATGAAGAGCCACTCCTAAGAAAGAGTTCTCGCCGGTTGTCTATCTTTCCAAT				
	140				
	250 AAGTCAAAATGAAGAGCCACTCCTAAGAAAGAGTTCTCGCCGGTTGTCTATCTTTCCAAT				
QY	141 CCAGTACCCCTGATATTTGGAAAAATGTATAAAACAGGCACAGGCTTCCTTCGGACAGCAGA				
	200				
	310 CCAGTACCCCTGATATTTGGAAAAATGTATAAAACAGGCACAGGCTTCCTTCGGACAGCAGA				
QY	201 AGAGTTGACTTATCAAAAGGATCTCCCTCACTCGGAACAAGCTTTAAAGCAGATGAGAAGTA				
	260				
	370 AGAGTTCGACTTATCAAAAGGATCTCCCTCACTCGGAACAAGCTTTAAAGCAGATGAGAAGTA				
DB	429				

```
QY 261 CTTTCATCTCTACATCTTAGCCCTTTTTCAGCCAGTGTGGAATTTGTAATGAAAATTTT 320
|
|
|
Db 430 CTTTCATCTCTACATCTTAGCCCTTTTTCAGCCAGTGTGGAATTTGTAATGAAAATTTT 489
|
|
|
QY 321 GGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTTCAGAGGCTCGCTGTTTCTATGGCTTTCA 380
|
|
|
Db 490 GGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTTCAGAGGCTCGCTGTTTCTATGGCTTTCA 549
|
|
|
QY 381 AATTCTCATCGAAGTGTTCACCTCAGAGATGTACAGTTTCTGTATAGACACTTACATCAG 440
|
|
|
Db 550 AATTCTCATCGAAGTGTTCACCTCAGAGATGTACAGTTTCTGTATAGACACTTACATCAG 609
|
|
|
QY 441 AGATCCCAAGAAAAGGAAATTTTATTTAATGCAATTTGAAACCAATGCCCTATGTTAAGAA 500
|
|
|
Db 610 AGATCCCAAGAAAAGGAAATTTTATTTAATGCAATTTGAAACCAATGCCCTATGTTAAGAA 669
|
|
|
QY 501 AAAAGCAGATTGGCCCTTCGGATGGATAGCAGATAGAAAATCTACTTTTGGGAAAAGAGT 560
|
|
|
Db 670 AAAAGCAGATTGGCCCTTCGGATGGATAGCAGATAGAAAATCTACTTTTGGGAAAAGAGT 729
|
|
|
QY 561 GGTGGCCTTTGCTGTAGAGGAGTTTCTTCTCAGGATCTTTTGTCTGCTATATTTCTG 620
|
|
|
Db 730 GGTGGCCTTTGCTGTAGAGGAGTTTCTTCTCAGGATCTTTTGTCTGCTATATTTCTG 789
|
|
|
QY 621 GCTTAAAGAGAGAGGCTTTATGCGCAGGACTCACTTTTCCAAATGAACTCATCAGCAGAGA 680
|
|
|
Db 790 GCTTAAAGAGAGAGGCTTTATGCGCAGGACTCACTTTTCCAAATGAACTCATCAGCAGAGA 849
|
|
|
QY 681 TGAAGGACTTCACTGTGACTTTTGTCTGCTGTGTTTCCAAATGAACTCATCAGCAGAGTTC 740
|
|
|
Db 850 TGAAGGACTTCACTGTGACTTTTGTCTGCTGTGTTTCCAAATGAACTCATCAGCAGAGTTC 909
|
|
|
QY 741 AGAAGAAAGGTCAGGAGATCATTTGATGCTGTCAAAATTTGAGCAGAGTTTCTTAAC 800
|
|
|
Db 910 AGAAGAAAGGTCAGGAGATCATTTGATGCTGTCAAAATTTGAGCAGAGTTTCTTAAC 969
|
|
|
QY 801 AGAAGCCTTGCAGGTTGGCCTCATTTGGAATGAATTCATTTGATGAAACAGTACATTGA 860
|
|
|
Db 970 AGAAGCCTTGCAGGTTGGCCTCATTTGGAATGAATTCATTTGATGAAACAGTACATTGA 1029
|
|
|
QY 861 GTTTGTAGCTGACAGATTACTTGTGAACTTGGAATCTCAAGGTTTTCAGGCGAGAAA 920
|
|
|
Db 1030 GTTTGTAGCTGACAGATTACTTGTGAACTTGGAATCTCAAGGTTTTCAGGCGAGAAA 1089
|
|
|
QY 921 TCTTTTGTATTTATGGAACAAATTTCTTTAGAGGAAACAAATTTCTTTTGGAGAACG 980
|
|
|
Db 1090 TCTTTTGTATTTATGGAACAAATTTCTTTAGAGGAAACAAATTTCTTTTGGAGAACG 1149
|
|
|
QY 981 AGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACACAGATAAAGTCTTACCTT 1040
|
|
|
Db 1150 AGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACACAGATAAAGTCTTACCTT 1209
|
|
|
QY 1041 GGATGCAGATTTT 1053
|
|
|
Db 1210 GGATGCAGATTTT 1222
|
|
|
```

RESULT 5

AY398973

LOCUS

AY398973 Homo sapiens HCM0069 gene, VIRUAL TRANSCRIPT, partial sequence, genomic survey sequence. 1056 bp DNA linear GSS 12-DEC-2003

ACCESSION

AY398973

VERSION

AY398973.1

KEYWORDS

GSS.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 1056)

AUTHORS

Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,

Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
 Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
 Science 302 (5652), 1960-1963 (2003)
 14671302
 2 (bases 1 to 1056)
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
 Direct Submission
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
 This sequence was made by sequencing genomic exons and ordering them based on alignment.
 Location/Qualifiers
 1..1056
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 <1..>1056
 /locus_tag="HCM0069"

Query Match 89.6%; Score 943.4; DB 9; Length 1056;
Best Local Similarity 89.6%; Pred. No. 6.2e-240;
Matches 944; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAAGCGCGGAGCGCGGGCTGGATCAGGATGAGAGATCATCTTCA 60
|
|
|
Db 1 ATGGGCGACCCGGAAAGCGCGGAGCGCGGGCTGGATCAGGATGAGAGATCATCTTCA 60
|
|
|

QY 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCG 120
|
|
|
Db 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCG 120
|
|
|

QY 121 CGGTTTGTCTATCTTCCAATCCAGTACCTGTATATTTGGAAAAATGATAAAGAGCAG 180
|
|
|
Db 121 CGGTTTGTCTATCTTCCAATCCAGTACCTGTATATTTGGAAAAATGATAAAGAGCAG 180
|
|
|

QY 181 GCTTCTCTCGGACGAGAGAGGTTGACTTATCAAGAGGATCTCCCTCAGTGGAAACAG 240
|
|
|
Db 181 GCTTCTCTCGGACGAGAGAGGTTGACTTATCAAGAGGATCTCCCTCAGTGGAAACAG 240
|
|
|

QY 241 CTTAAAGCAGATGAGAGTACTTTCATCTCTCACATCTTAGCCTTTTTCGAGCCAGTGT 300
|
|
|
Db 241 CTTAAAGCAGATGAGAGTACTTTCATCTCTCACATCTTAGCCTTTTTCGAGCCAGTGT 300
|
|
|

QY 301 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCT 360
|
|
|
Db 301 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCT 360
|
|
|

QY 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTG 420
|
|
|
Db 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTG 420
|
|
|

QY 421 CTGATGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTATTTAATGCAATTGAA 480
|
|
|
Db 421 CTGATGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTATTTAATGCAATTGAA 480
|
|
|

QY 481 ACCATGCCCTATGTTAAGAAAAGAGATTTGGGCTTTCGATGGATAGCAGATAGAAAA 540
|
|
|
Db 481 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 540
|
|
|

QY 541 TCTACTTTTGGGAAAAGAGTGGTGGCTTTGCTGCTAGAGAGGTTTCTTCTCAGGA 600
|
|
|
Db 541 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 600
|
|
|

QY 601 TCTTTTGTCTATATTTCTGGCTAAAAGAGAGGTTCTTATGCCAGGAGTCACTTTTTTCC 660
|
|
|
Db 601 TCTTTTGTCTATATTTCTGGCTAAAAGAGAGGTTCTTATGCCAGGAGTCACTTTTTTCC 660
|
|
|

```
QY 661 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db |||||
QY 721 TACTTAGTAATAAGCCTTCAGAAGAAAGGTCAGGAGATCAATTGATGCTGTCAA 780
Db |||||
QY 781 ATTGACGACGAGTTTAAACAGAGACCTTGCCAGTTGGCCTCATTCGGAATGAATTCGATT 840
Db |||||
QY 841 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTTGGAACCTTGGAATCTCA 900
Db |||||
QY 901 AGGTTTTTCAGGCGAGAAATCCTTTTGATTTATGGAACAACTTTCTTTAGAGGAA 960
Db |||||
QY 961 ACAAAATTTCTTGAGAAACGAGTTTCAGAGTATCAGGTTTTCAGTTATGGCAGAAACC 1020
Db |||||
QY 1021 ACAGATAAGCTCTTCACTTGGATGAGATTTT 1053
Db |||||
QY 1021 ACAGATAAGCTCTTCACTTGGATGAGATTTT 1053
Db |||||
```

RESULT 6

```
AL547501
LOCUS 1083 bp mRNA linear EST 25-MAR-2004
DEFINITION AL547501 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1011YF14 5-PRIME, mRNA sequence.
ACCESSION AL547501
VERSION AL547501.3 GI:45747948
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:31269332.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4436.r
```

For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0D1011DC070P1&c=4436.r.

FEATURES

```
source
1..1083
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1011YF14"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
```

ORIGIN

Query Match

88.1%; Score 928; DB 1; Length 1083;

RESULT 7

```
Best Local Similarity 97.9%; Pred. No. 7.9e-236;
Matches 975; Conservative 6; Mismatches 11; Indels 4; Gaps 4;
QY 1 ATGGGCGACCCGAAAGCGCGGCGG-CCTGGATCAGATCAGAGATCATCTTC 59
Db 27 ATGGGCGACCCGAAAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 86
QY 60 AGACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCG 119
Db 87 AGACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCG 146
QY 120 CCGGTTTGTTCATCTTCCAAATCCAGTACCCTGATATTTGAAAAATGATATAACAGGCACA 179
Db 147 CCGGTTTGTTCATCTTCCAAATCCAGTACCCTGATATTTGAAAAATGATATAACAGGCACA 206
QY 180 GGCCTTCCTTCGACAGCAGAGAGGTTGACTTATCAAGAGATCCCTCACTGGAACAA 239
Db 207 GGCCTTCCTTCGACAGCAGAGAGGTCGACTTATCAAGAGATCCCTCACTGGAACAA 266
QY 240 GCTTAAAGCAGATGAGAAAGTACTTCTCTCATCTTACATCTTAGCCCTTTTTCAGCCAGTGA 299
Db 267 GCTTAAAGCAGATGAGAAAGTACTTCTCTCATCTTACATCTTAGCCCTTTTTCAGCCAGTGA 326
QY 300 TGGAAATTTGTAATGAAAAATTTGCTGGAGCCCTTTAGTCAGAGGTCAGGTTCCAGAGGC 359
Db 327 TGGAAATTTGTAATGAAAAATTTGCTGGAGCCCTTTAGTCAGAGGTCAGGTTCCAGAGGC 386
QY 360 TCGCTGTTCTATGGCTTTCAAAATTCATCGAGAATGTTCACTCAGAGATGTCAGATTT 419
Db 387 TCGCTGTTCTATGGCTTTCAAAATTCATCGAGAATGTTCACTCAGAGATGTCAGATTT 446
QY 420 GCTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTTGA 479
Db 447 GCTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTTGA 506
QY 480 AACCATGCCCTATGTTAAAGAAAAAGCAGATTTGGGCCCTTGGATGGATAGCAGATAGAAA 539
Db 507 AACCATGCCCTATGTTAAAGAAAAAGCAGATTTGGGCCCTTGGATGGATAGCAGATAGAAA 566
QY 540 ATCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAGAGGATTTTCTCTCAGG 599
Db 567 ATCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAGAGGATTTTCTCTCAGG 626
QY 600 ATCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGCTTATGCCAGGACTCAGCTTTTC 659
Db 627 ATCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGCTTATGCCAGGACTCAGCTTTTC 686
QY 660 CAATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTCTTGCCTGATGTTCCA 719
Db 687 CAATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTCTTGCCTGATGTTCCA 746
QY 720 ATACTTAGTAATAAGCCTTCAGAAGAAAGGTCAGGAGATCAATTGTTGATGCTGTCAA 779
Db 747 ATACTTAGTAATAAGCCTTCAGAAGAAAGGTCAGGAGATCAATTGTTGATGCTGTCAA 806
QY 780 AATTGACGAGGCTTTTAAACAGAGCCTTGCAGGTTGGCCTCAATGGAATGAATTCGAT 839
Db 807 AATTGACGAGGCTTTTAAACAGAGCCTTGCAGGTTGGCCTCAATGGAATGAATTCGAT 866
QY 840 TTTGATGAACACAGTACATTCAGTTTGTAGCTGACAGATTAATCTTG-TGGAACTTGGAATTC 898
Db 867 TTTGATGAACACAGTACATTCAGTTTGTAGCTGACAGATTAATCTTG-TGGAACTTGGAATTC 926
QY 899 CAAAGGTTTTTCAGGCGAGAAAAATTCCTTTTGAATTTTATGAAAAACATTTCTTTTAGAGAA 958
Db 927 CAAAGGTTTTTCAGGCGAGAAAAATTCCTTTTGAATTTTATGG-AAAMAATTTCTTTTAGAAGG-A 984
QY 959 AAACAAATTTCTTTGAGAAACGAGTTTCAGAGATTC 994
Db 985 AAAMAAATTTCTTTKAGAACGAGTTTMRAGMWCAGC 1020
```

```
AY398974      1008 bp      DNA      linear      GSS 12-DEC-2003
LOCUS      Pan troglodytes HWM0069 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION      genomic survey sequence.
ACCESSION      AY398974
VERSION      AY398974.1 GI:39754963
KEYWORDS      GSS.
SOURCE      Pan troglodytes (chimpanzee)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
REFERENCE      1 (bases 1 to 1008)
AUTHORS      Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
      Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
      Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sniinsky,J.J.,
      Adams,M.D. and Cargill,M.
TITLE      Direct Submission
COMMENT      Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
      Rockville, MD 20850, USA
      This sequence was made by sequencing genomic exons and ordering
      them based on alignment.
FEATURES             Location/Qualifiers
     source            1..1008
                        /organism="Pan troglodytes"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9598"
                        <1..>1008
                        /locus_tag="HCM0069"
ORIGIN
Query Match      78.7%; Score 828.2; DB 9; Length 1008;
Best Local Similarity 82.6%; Pred. No. 3e-209;
Matches 830; Conservative 0; Mismatches 175; Indels 0; Gaps 0;
QY 49 AGATCATCTTCAGACACCAACGAAAGTGAATTAAGTCAAACTCAATGAAGAGCCACTCTTAAGA 108
Db |||||
QY 109 AAGAGTTCTCGCGGTTTGTCTCTTCCAAATCCAGTACCCTGATATTTGGAAATGTAT 168
Db |||||
QY 169 AAACAGGCACAGGCTTCCTTCTGGACAGAGAAGAGTTGACTTTATCAAGGATCTCCCT 228
Db |||||
QY 229 CACTGGAACAAGCTTAAAGCAGATGAGAAGTACTTCATCTCTCACATCTTAGCCTTTTT 288
Db |||||
QY 181 CACTNGAACACGCTTAAAGCAGATGAGAAGTACTTCATCTCTCACATCTTAGCCTTTTT 240
QY 289 GCAGCCAGTGTGGAAATGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAG 348
Db |||||
QY 241 GCAGCCAGTGTGGAAATGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAG 300
QY 349 GTTCAGAGGCTCGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAGAATGTTCACTCAGAG 408
Db |||||
QY 301 GTTCAGAGGCTCGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAGAATGTTCACTCAGAG 360
QY 409 ATGTACAGTTTCTCATGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTT 468
Db |||||
QY 361 ATGTACAGTTTCTCATGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTT 420
QY 469 AATGCAATTGAAACCATGCCCTATGTTAAGAAAGGAGCAGATGGCGCTTGGCATGATA 528
```

```
Db 421 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 480
QY 529 GCAGATAGAAAATCTACTTTTGGGGAAGAGTGGTGGCCCTTGGCTGCTGTAGAAGGAGTT 588
Db |||||
QY 481 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 540
QY 589 TTCTTCTCAGGATCTTTTGTCTGTATATTCTGGCTAAAGAGAGAGGTCTTATGCGCAGGA 648
Db |||||
QY 541 TTCTTCTCAGGATCTTTTGTCTGTATATTCTGGCTAAAGAGAGAGGTCTTATGCGCAGGA 600
QY 649 CTCACATTTTCCAAATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGC 708
Db |||||
QY 601 NTCACATTTTCCAAATGAACATCANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 660
QY 709 CTGATGTTCCAAATGACTTAGTAAATAGCCCTTCAGAAAGAGGTTCAGGAGATCATTTGTT 768
Db |||||
QY 661 CTGATGTTCCAAATGACTTAGTAAATAGCCCTTCAGAAAGAGGTTCAGGAGATCATTTGTT 720
QY 769 GATGCTGTCAAAATTTGAGCAGGAGGATTTTAAACAGAAGCCTTGGCCTCATTTGGA 828
Db |||||
QY 721 GATGCTGTCAAAATTTGAGCAGGAGGATTTTAAACAGAAGCCTTGGCCTCATTTGGA 780
QY 829 ATCAATTTGCAATTTTGTATGAACAGTACATTTGAGTTTGTAGCTGACAGATTAATTTGTGAA 888
Db |||||
QY 781 ATCAATTTGCAATTTTGTATGAACAGTACATTTGAGTTTGTAGCTGACAGATTAATTTGTGAA 840
QY 889 CTTGGAATTTCAAAAGGTTTTCAGGCAGAAAATCCCTTTGATTTATTTGGAAGAAATTTCT 948
Db |||||
QY 841 CTTGGAATTTCAAAAGGTTTTCAGGCAGAAAATCCCTTTGATTTATTTGGAAGAAATTTCT 900
QY 949 TTAGAAGGAAAACAAATTTCTTTTGAGAAACGAGTTTCAGAGTATCATGCGTTTTCGAGTT 1008
Db |||||
QY 901 TTAGAAGGAAAACAAATTTCTTTTGAGAAACGAGTTTCAGAGTATCATGCGTTTTCGAGTT 960
QY 1009 ATGCGAGAAACACAGATTAACGCTTTCACCTTGGATGAGATTTT 1053
Db |||||
QY 961 ATGCGAGAAACACAGATTAACGCTTTCACCTTGGATGAGATTTT 1005
```

```
RESULT 8
LOCUS      BQ441857
DEFINITION      AGENCOURT_7827448 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:6014179
      S', mRNA sequence.
ACCESSION      BQ441857
VERSION      BQ441857.1 GI:21180933
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 879)
AUTHORS      NIH-MGC http://mgi.nci.nih.gov/
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
      Email: cgabbs@mail.nih.gov
      Tissue Procurement: ATCC
      cDNA Library Preparation: Life Technologies, Inc.
      DNA Sequencing by: Agencourt Bioscience Corporation
      Clone distribution: MGC clone distribution information can be
      found through the I.M.A.G.E. Consortium/LLNL at:
      http://image.llnl.gov
      Plate: LLAM13208 row: d column: 20
      High quality sequence stop: 662.
      Location/Qualifiers
          1..879
          /organism="Homo sapiens"
          /mol_type="mRNA"
          /db_xref="taxon:9606"
          /clone="IMAGE:6014179"
          /tissue_type="embryonal carcinoma, cell line"
```

```
/lab host="DH10B (phage-resistant)"
/clone lib="NIH_MGC_92"
/organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 Kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 75.2%; Score 791.6; DB 5; Length 879;
Best Local Similarity 98.6%; Pred. No. 1.6e-199;
Matches 820; Conservative 0; Mismatches 9; Indels 3; Gaps 2;

QY 1 ATGGCGACCCGGAAGGCGGCGGCTGGATCAGATGAGAGATCATCTTCA 60
DB 22 ATGGCGACCCGGAAGGCGGCGGCTGGATCAGATGAGAGATCATCTTCA 81

QY 61 GACACCAAGAAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAAGATTTCTGC 120
DB 82 GACACCAAGAAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAAGATTTCTGC 141

QY 121 CGTTTGTGATCTTCCAAATCAGTACCTGATATTTGAAATGTATTAACAGGCACAG 180
DB 142 CGTTTGTGATCTTCCAAATCAGTACCTGATATTTGAAATGTATTAACAGGCACAG 201

QY 181 GCTTCTCTTGACAGCAGAGGTGACTTATCAAGAGTCTCCCTCACTGGAACAAG 240
DB 202 GCTTCTCTTGACAGCAGAGGTGACTTATCAAGAGTCTCCCTCACTGGAACAAG 261

QY 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTTAGCCCTTTTTCAGCCAGTGTAT 300
DB 262 CTTAAAGCAGATGAGAAGTACTTCTCATCTTAGCCCTTTTTCAGCCAGTGTAT 321

QY 301 GGAATTGTAATGAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCGAGTTCAGAGGCT 360
DB 322 GGAATTGTAATGAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCGAGTTCAGAGGCT 381

QY 361 CGCTGTTCTATGGCTTTCAAATCTCATCGAGATGTTCACTCAGAGATGTACAGTTTG 420
DB 382 CGCTGTTCTATGGCTTTCAAATCTCATCGAGATGTTCACTCAGAGATGTACAGTTTG 441

QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
DB 442 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 501

QY 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGGCTTCGCGATGATAGATGAAAA 540
DB 502 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGGCTTCGCGATGATAGATGAAAA 561

QY 541 TCTACTTTTGGGAAAGTGGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 600
DB 562 TCTACTTTTGGGAAAGTGGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 621

QY 601 TCTTTTGTGCTATATTCGGCTTAAAGAGAGAGGTCTTTATGCCAGGACTCACTTTTTC 660
DB 622 TCTTTTGTGCTATATTCGGCTTAAAGAGAGAGGTCTTTATGCCAGGACTCACTTTTTC 681

QY 661 AATGAACTCATCAGCAGATGAAGGCTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
DB 682 AATGAACTCATCAGCAGATGAAGGCTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 741

QY 721 TACTTAGTAAATTAAGCCTTCAGAAAGAGG--TCAGGAGATCATTTGTGCTGTCA 778
DB 742 TACTTAGTAAATTAAGCCTTCAGAAAGAGGCTTCAGGAGATCATTTGTGCTGTCA 801

QY 779 AATT- GAGCAGAGTTTTTAACAGAGCCTTGCCAGTTGGCTCAATCGAA 829
DB 802 AATTGGAGCAGAGTTTTTAACAGAGCCTTGCCAGTTGGCTCAATTTGGA 853
```

```
RESULT 9
AY398975 1043 bp DNA linear GSS 12-DEC-2003
LOCUS
```

```
DEFINITION Mus musculus HCM0069 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY398975
VERSION AY398975.1 GI:39754964
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1043)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1043)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..1043
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<J..>1043
/locus_tag="HCM0069"
gene
ORIGIN
Query Match 73.9%; Score 778.6; DB 9; Length 1043;
Best Local Similarity 81.3%; Pred. No. 5e-196;
Matches 855; Conservative 0; Mismatches 184; Indels 13; Gaps 1;

QY 1 ATGGGCGACCCCGAAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
DB 1 ATGGGCGACCCCGAAAGCGCGGAGCGCGCAGCGGAGAGGTGAGCAATTGTGTTCC 60

QY 61 GACACCAAGAAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAAGATTTCTCGC 120
DB 61 GACACCAAGAAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAAGATTTCTCGCA 120

QY 121 CGGTTTGTCTATCTTTCCAAATCCAGTACCTGATATTTGGAAATGTATAACAGGCACAG 180
DB 121 CGGTTTGTCTATCTTTCCAAATCCAGTATCTCTGAGATGTATCAAGCAAGACAG 180

QY 181 GCTTCTCTTCTGGACAGCAGAGAGGTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 240
DB 181 GCATCTTTCTGGACAGCAGAGAGGTGACTTGTCAAGAGACCTCCCTCACTGGAACAAG 240

QY 241 CTTAAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTAGCCCTTTTTCAGCCAGTGTAT 300
DB 241 CTTAAATCCGACGAGAAGTATTTTATCTCCACATCTTAGCCCTTTTTCAGCCAGTGTAT 300

QY 301 GGAATTGTAATGAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCGAGGTTCAGAGGCT 360
DB 301 GGAATTGTAATGAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCGAGGTTCAGAGGCT 360

QY 361 CGCTGTTCTATGGCTTTTCAAAATTTCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG 420
DB 361 CGCTGTTCTATGGCTTTTCAAGATTTCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTA 420

QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
DB 421 CTAATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
```

```
QY 481 ACCATGCCCTATCTTAAGAAAAAAGCAGATTGGCCCTTGCATGGATGATGAGATGAGAAA 540
Db NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTGTCTGCTGTGAGAGAGTGTTCCTCTCAGGA 600
Db NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
QY 601 TCTTTTGTCTCTATATTCTGGCTTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTCC 660
Db TCAATTTGTGCAATATTCTGGCTTAAAGAGAGAGGCTCATGCTGAGCTGACTTTTCA 660
QY 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTCTGCTGTGATGTTCCAA 720
Db AATGAACCTCATCAGCAGGAGTGAAGGACTTCACTGTGACTTTTCTGCTGTGATGTTCCAG 720
QY 721 TACTTAGTAAATGAAGCTTTCAGAAAGAGGTCAGGAGATCAATGTTGATGCTGTCAAA 780
Db TACTTGTGTAACAAGCTTTCAGAAAGATAGAGTGAAGGAAATCAATGCTGATGCTGTCAA 780
QY 781 ATTGACAGGAGTTTTAAACAGAGGCTTGCACAGTGGCTCAATGGAAATGAATTCGATT 840
Db ATCGAG-----CAGAAGCCTTGGCTTGGCTTCAATGGAAATGAATTCGATT 840
QY 841 TTGATGAAACAGTACATTCAGTGTGTAGCTGACAGATTAATCTGTGGAACCTTGGATTCCTCA 900
Db TTGATGAAACAGTACATTCAGTGTGTAGCTGACAGATTAATCTGTGGAACCTTGGATTCCTCA 887
QY 901 AAGTTTTCAGCAGAAATCCTTTTGTATTTATGAAACAACTTTCTTTAGAGGAA 960
Db AAGATTTTCAGCAGAAATCCTTTTGTATTTATGAGAAACATTTTCATTAGAAGGAA 947
QY 961 ACAATTTCTTTCAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020
Db ACAATTTCTTTCAGAAACAGTTTCTGAGTATCAGCGATTTTCAGTATGCGCCGAAACC 1007
QY 1021 ACAGATAACGCTTTCACCTTGGATGCAGATTT 1052
Db ACAGATAATGCTTTCACCTTGGATGCAGATTT 1039
```

```
RESULT 10
CN163214/c
LOCUS CN163214 892 bp mRNA linear EST 02-APR-2004
DEFINITION 952895 MARC 4PIG Sus scrofa cDNA 3', mRNA sequence.
ACCESSION CN163214
VERSION CN163214.1 GI:46177644
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 892)
AUTHORS Smith,T.P.L., Freking,B.A., Ford,J.J., Vallet,J.L., Wise,T.A.,
Noneman,D.J., Wray,J.E. and Keele,J.W.
TITLE Porcine EST collection using a normalized library constructed from
embryos representing early developmental stages
JOURNAL Unpublished (2003)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross_match v0.990329.
Plate: TMW8060 row: G column: 4
Seq primer: TAGAAGGCACATCGCAGG.
1..892 Location/Qualifiers
/organism="Sus scrofa"
/mol_type="mRNA"
```

```
ORIGIN
Query Match 69.4%; Score 731.2; DB 7; Length 892;
Best Local Similarity 94.1%; Pred. No. 2e-183;
Matches 760; Conservative 0; Mismatches 48; Indels 0; Gaps 0;
QY 246 AGCAGATGAGAAGTACTTCTCATCTCTACATCTTAGCCCTTTTTCAGCCACTGATGGAA 305
Db 892 ATCAGATGAGAAGTATTTTATCTCTCATCTCTTAGCCCTTTTTCAGCCACTGATGGAA 833
QY 306 TGTAAATGAAAAATTTTGGTGGAGCGCTTTTAGTCAGGAGTGCAGGTTCCAGAGGTCGCTG 365
Db 832 TGTCAATGAAAACTTGGTGGAGCGCTTTTAGTCAGGAGTGCAGGTTCCAGAGGTCGCTG 773
QY 366 TTTCTATGGCTTTCAAATTCATTCGAGAATGTTCTACTCAGAGATGATGATGTTGCTGAT 425
Db 772 TTTCTACGGCTTTCAAATTCATTCGAGAATGTTCTACTCAGAGATGATGATGTTGCTGAT 713
QY 426 AGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATGAAACCAT 485
Db 712 AGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATGAGACAA 653
QY 486 GCCTTATGTTAAGAAAAAAGCAGATTCGGCTTTCGATGATGATGATGATGATGATGATGAT 545
Db 652 GCATATGTTAAGAAAAAAGCAGATTCGGCTTTCGATGATGATGATGATGATGATGATGAT 593
QY 546 TTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAAAGAGTGTTCCTCAGGATCTTT 605
Db 592 TTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAAAGAGTGTTCCTCAGGATCTTT 533
QY 606 TGTCTGTATATTCGCTTAAAGAGAGAGTCTTATGCGAGGACTCACTTTTCCAAATGA 665
Db 532 TGTCTGTATATTCGCTTAAAGAGAGAGTCTTATGCGCTGAGTCACTTTTCCAAATGA 473
QY 666 ACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGCCTGATGTTTCCAAATCACT 725
Db 472 ACTCATCAGCAGAGATGAAGGAGTCTTCACTGTGACTTTTGCCTGATGTTTCCAAATCACT 413
QY 726 AGTAAATAGCCCTTCAGAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCAAATGA 785
Db 412 GGTAAATAGCCCTTCAGAGCAAGAGTTCAGGAGATCATTTGTTGATGCTGTCAAATGA 353
QY 786 GCAGGAGTTTAAACAGAGCCCTTGCAGTTGGCCTCATTTGGAATGAATTTGATTTGAT 845
Db 352 GCAGGAGTTTAAACAGAGCCCTTGCAGTTGGCCTCATTTGGAATGAATTTGATTTGAT 293
QY 846 GAAACAGTACATTTGAGTTTGTAGCTGACAGATTAATGTTGSACTTGGATTTCTCAAGGT 905
Db 292 GAAACAGTACATTTGAGTTTGTAGCTGACAGATTAATGTTGSACTTGGATTTCTCAAGGT 233
QY 906 TTTTTCAGGAGAAAAATCCTTTTTCATTTTATGAAAAACATTTTCTTGAAGAAAAACAAA 965
Db 232 TTTTTCAGGAGAAAAATCCTTTTTCATTTTATGAAAAACATTTTCTTGAAGAAAAACAAA 173
QY 966 TTTCTTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGAGTTATGCGAGAAACACAGA 1025
Db 172 TTTCTTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGAGTTATGCGAGAAACACAGA 113
QY 1026 TAACGTCCTTACCTTCGATGTCAGATTTT 1053
Db 112 TAATGTCCTTACCTTCGATGTCAGATTTT 85
```

RESULT 11
CF995079
LOCUS

CF995079 914 bp mRNA linear EST 25-NOV-2003


```

DEFINITION      AGENCOURT_15621478 NIH_MGC_147 Homo sapiens cDNA clone
IMAGE:30528806 5', mRNA sequence.
ACCESSION       CF995079
VERSION         CF995079.1 GI:38511139
KEYWORDS        EST.
SOURCE          Homo sapiens (human)
ORGANISM        Homo sapiens
REFERENCE       1 (bases 1 to 914)
AUTHORS        Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE          Eukaryota; Mammalia; Primates; Catarrhini; Hominidae; Homo.
JOURNAL        NIH-MGC http://mgi.nci.nih.gov/.
COMMENT        National Institutes of Health, Mammalian Gene Collection (MGC)
                Unpublished (1999)
                Contact: Daniela S. Gerhard, Ph.D.
                Office of Cancer Genomics
                National Cancer Institute / NIH
                Bldg. 31 Rm10A07 Bethesda, MD 20892
                Email: cgabbs@mail.nih.gov
                Tissue procurement: Dr. Stefan Hansson
                and advice from Piero Carninci (RIKEN)
                CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                DNA Sequencing by: Agencourt Bioscience Corporation
                Clone distribution: MGC clone distribution information can be
                found through the I.M.A.G.E. Consortium/LLNL at:
                http://image.llnl.gov
                Plate: NDAM615 row: f column: 15
                High quality sequence stop: 677.
FEATURES        Location/Qualifiers
                 1..914
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:30528806"
                    /tissue_type="Human Placenta"
                    /lab_host="DH10B TONa"
                    /clone_lib="NIH MGC 147"
                    /note="Organ: Placenta; Vector: pBluescriptR; Site: 1:
                    all-XhoI, Site 2: BamH; Oligo-dT primed using primer
                    5'-TTTTTTTTTTTTTTVN-3', size-selected for average
                    insert size 2.3 kb and normalized to ROT 5. This is a
                    primary library enriched for full-length clones and
                    constructed using the Cap-trapper method (Carninci, in
                    preparation). Library constructed by M. Brownstein
                    (NIH/NHGRI, National Institutes of Health). Note: This is
                    a NIH_MGC library."
ORIGIN
Query Match      69.0%; Score 726.6; DB 7; Length 914;
Best Local Similarity 99.0%; Pred. No. 3.4e-182;
Matches 763; Conservative 0; Mismatches 4; Indels 4; Gaps 3;

QY 1 ATGGGCGACCCGGAAGCGCGGCGGCTGGATCAGGATGAGATCATCTTCA 60
    |||
DB 87 ATGGGCGACCCGGAAGCGCGGCGGCTGGATCAGGATGAGATCATCTTCA 146
    |||

QY 61 GACACCAAGCAAGTCAATTAAGTCAATGAAGACCCACTCTTAAGAAAGATTCTCG 120
    |||
DB 147 GACACCAAGCAAGTCAATTAAGTCAATGAAGACCCACTCTTAAGAAAGATTCTCG 206
    |||

QY 121 CGGTTTGTCATCTTTCCATCCAGTACCTGTATTTGGAAATGTATTAACAGGCACAG 180
    |||
DB 207 CGGTTTGTCATCTTTCCATCCAGTACCTGTATTTGGAAATGTATTAACAGGCACAG 266
    |||

QY 181 GCTTCCTCTCGACAGCAGAGGTTGACTTATCAAGGATCTCCCTCACTCGAACAAAG 240
    |||
DB 267 GCTTCCTCTCGACAGCAGAGGTTGACTTATCAAGGATCTCCCTCACTCGAACAAAG 326
    |||

QY 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTTCAATCTTAGCCCTTTTTCAGGCATGAT 300
    |||
DB 327 CTTAAAGCAGATGAGAAGTACTTCTCATCTTCAATCTTAGCCCTTTTTCAGGCATGAT 386
    |||

QY 301 GGAATTGTAATAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCGAGGCTTCCAGGCT 360
    |||

```

```

Db 387 GGAATTGTAATAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCGAGGTTCCAGAGCT 446
    |||
QY 361 CGCTGTTTCTATGGCTTTCATAATCTCATCGAGATGTTCACTCAGAGATGTACAGTTG 420
    |||
DB 447 CGCTGTTTCTATGGCTTTCATAATCTCATCGAGATGTTCACTCAGAGATGTACAGTTG 506
    |||
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAATTTTATTAATGCAATTGAA 480
    |||
DB 507 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAATTTTATTAATGCAATTGAA 566
    |||
QY 481 ACCATCCCTATCTTAAGAAAAAGCAGATTGGGCCCTTGCATGGATAGCAGATAGAAA 540
    |||
DB 567 ACCATCCCTATCTTAAG--AAAAGCAGATTGGGCCCTTGCATGGATAGCAGATAGAAA 624
    |||
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGATTTTCTTCTCAGGA 600
    |||
DB 625 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGATTTTCTTCTCAGGA 684
    |||
QY 601 TCTTTTGTGCTATATTTCTGGCTAAAGACAGAGAGTCTTATGCCAGGACTCAGTTTTTC 660
    |||
DB 685 TCTTTTGTGCTATATTTCTGGCTAAAGACAGAGAGTCTTATGCCAGGACTCAGTTTTTC 744
    |||
QY 661 AATGAACACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA 720
    |||
DB 745 AATGAACACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA 804
    |||
QY 721 TACTTAGTAAATAAGCC--TTCAGAAAGAGGTC-AGGAGATCATTTGTTG 769
    |||
DB 805 TACTTAGTAAATAAGCCTTTCAAAGAAAGGGTCTAGGAAGATCATTTGTTG 855
    |||

RESULT 12
LOCUS      CO648380
DEFINITION ILLUMIGEN MCQ 41729 Katze MMPB2 Macaca mulatta cDNA clone
            IBIUM:25149 5' similar to Bases 5 to 770 highly similar to human
            RRM2B (Hs.512592), mRNA sequence.
ACCESSION  CO648380
VERSION     CO648380.1 GI:50569874
KEYWORDS    EST.
SOURCE      Macaca mulatta (rhesus monkey)
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
            Cercopitheinae; Macaca.
REFERENCE   1 (bases 1 to 870)
AUTHORS     Katze,M.G., Thomas,M., Korth,M., Iadonato,S.P. and Magness,C.L.
TITLE       Large-scale Rhesus Macaque cDNA Sequencing
JOURNAL     Unpublished (2003)
COMMENT     Contact: C. Magness
            Illumigen Biosciences Inc.
            2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
            Tel: 2063780400
            Fax: 2063780408
            Email: cmagness@illumigen.com
            Sequenced on 2004.06.02. 691 Q20 bases. Library Preparation: Prof.
            Michael Katze lab at University of Washington DNA Sequencing:
            Illumigen Biosciences Inc. For further information, see
            http://www.macaque.org
PCR Primers
FORWARD: CCTCACTAAGGGAACAAA
BACKWARD: CACTATAGGCGAATTGGTA
Insert Length: 870 Std Error: 0.00
Plate: CL000348 row: B column: 06
Seq primer: CCCTCACTAAGGGAACAAA
POLYA=No.
FEATURES    Location/Qualifiers
             1..870
                 /organism="Macaca mulatta"
                 /mol_type="mRNA"
                 /strain="Indian"
                 /db_xref="taxon:9544"

```



```
/clone="IBIUW:25149"
/sex="male"
/tissue_type="blood"
/cell_type="PBMC"
/dev_stage="adult"
/lab_host="Electromax DH10B"
/clone_lib="Katze MRPB2"
/notes="Vector: pDONR 222; Site 1: BarG I; Site 2: BarG I;
Created from Cloneminer cDNA Library Construction kit
(catalog #18249-029)"

ORIGIN
Query Match 67.5%; Score 710.6; DB 7; Length 870;
Best Local Similarity 98.8%; Pred. No. 6.2e-178;
Matches 716; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 329 GCTTTAGTCAGAGGTGCAGGTTCCAGAGCGTCGCTGTTCTATGGCTTTCAAATTCCTCA 388
Db |
QY 389 TCGAGAACTTCACTCAGAGGTGTACAGTTTCTGTATACACACTTACATCAGAGATCCCA 448
Db |
QY 63 TCGAGAACTTCACTCAGAGGTGTACAGTTTCTGTATACACACTTACATCAGAGATCCCA 122
QY 449 AGAAAAGGGAATTTTATTAATGCAATTGAAACCAATGCCCTATGTTAAGAAAAGAGCAG 508
Db |
QY 123 AGAAAAGGGAATTTTATTAATGCAATTGAAACCAATGCCCTATGTTAAGAAAAGAGCAG 182
QY 509 ATTGGGCTTGGATGGATGAGATAGAGAAATCTACTTTTGGGAAAGAGTGGTGGCT 568
Db |
QY 183 ATTGGGCTTGGATGGATGAGATAGAGAAATCTACTTTTGGGAAAGAGTGGTGGCT 242
QY 569 TTGCTGCTGAGAGAGTTTCTTCTCAGAGATCTTTGCTGCTATATTTCTGGCTAAAGA 628
Db |
QY 243 TTGCTGCTGAGAGAGTTTCTTCTCAGAGATCTTTGCTGCTATATTTCTGGCTAAAGA 302
QY 629 ACAGAGGCTTATGCGAGGACTCCTTTTCCAAATGAATCTATCAGAGATGAGGAC 688
Db |
QY 303 ACAGAGGCTTATGCGAGGACTCCTTTTCCAAATGAATCTATCAGAGATGAGGAC 362
QY 689 TTCACTGTGACTTTGCTTGGCTGATGTTCCAAATGATAGTAATAGCCTTCAGAGAAA 748
Db |
QY 363 TTCACTGTGACTTTGCTTGGCTGATGTTCCAAATGATAGTAATAGCCTTCAGAGAAA 422
QY 749 GGGTCAGGAGATCAATTTGTGATGCTGCAAAATTTAGCAGAGATTTTAAACAGAGCT 808
Db |
QY 423 GGGTCAGGAGATCAATTTGTGATGCTGCTCAAAATTTAGCAGAGATTTTAAACAGAGCT 482
QY 809 TGCCAGTTGGCTCATTTGGAATGCAATTTGATGAAACAGTACATTTGATTTGTAG 868
Db |
QY 483 TGCCAGTTGGCTCATTTGGAATGCAATTTGATGAAACAGTACATTTGATTTGTAG 542
QY 869 CTGACAGATTACTTTGGAATTTGATTTCTCAAAAGTTTTCAGGAGAGAAATTCCTTTTG 928
Db |
QY 543 CTGACAGATTACTTTGGAATTTGATTTCTCAAAAGTTTTCAGGAGAGAAATTCCTTTTG 602
QY 929 ATTTATGGAACAACTTTCTTTAGAGGAGAAACAAATTTCTTTGAGAAACAGTTTCAG 988
Db |
QY 603 ATTTATGGAACAACTTTCTTTAGAGGAGAAACAAATTTCTTTGAGAAACAGTTTCAG 662
QY 989 AGTATCAGGTTTTCAGTTATGGCAGAAACACAGATTAACCTTTCCCTTGGATGTCAG 1048
Db |
QY 663 AGTATCAGGTTTTCAGTTATGGCAGAAACACAGATTAACCTTTCCCTTGGATGTCAG 722
QY 1049 ATTTT 1053
Db |
QY 723 ATTTT 727
```

```
RESULT 13
CD656661
LOCUS CD656661 804 bp mRNA linear EST 18-JUN-2003
DEFINITION AGENCOURT_14555875 NIA Human H1 Embryonic Stem Cell cDNA Library
```

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

(Long) Homo sapiens cDNA clone IMAGE:30427288 5', mRNA sequence.
CD656661 GI:31897180
EST.
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 804)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabbs@mail.nih.gov
Tissue Procurement: Irene Ginis and Mahendra Rao, NIA
cDNA Library Preparation: Yulan Piao and Minoru KO
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC cDNA Library
can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: NDAM514 row: h column: 17
High quality sequence stop: 680.

FEATURES
source

1..804
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30427288"
/tissue_type="Embryonic Stem cells"
/cell_line="WA01"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIA Human H1 Embryonic Stem Cell cDNA Library (Long)"
/notes="Vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI;
This is a long-transcript enriched cDNA library (Genome
Res. 11: 1553-1558 (2001). [PMID: 11544199]) from WA01
cell line. Undifferentiated human ES cell line WA01/H1
was obtained from WiCell Research Institute, Inc.
Madison, WI, cultured according to their instructions, on
MBF feeders. They formed round colonies with defined edges
and were positive for alkaline phosphatase, SSEA-4, OCT3,
OCT4, REX1, UTF, TERT, SOX2, CX43 and CX45. They are
negative for GATA2, GATA4, PDX1, NCAM, MSX1, FUT3, SSEA-1,
TUBB3, NES, GFAP, and BOMES. When confluent (18-10 days
after plating), the ES cells from 4 x 6cm dishes were
treated with 1 mg/ml collagenase, type IV
(Invitrogen/GIBCO) for 5-10 min and gently scraped off
with 5 ml pipette. RNA was purified with Trizol Reagent
from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558
(2001). [PMID: 11544199]) Double-stranded cDNAs were
synthesized with an Oligo(dT) primer [Invitrogen:
5'-pGACTAGTTCAGTCGCGCGCCCTCTTTTCTTTTCTTTT-3'] from
3.4g of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to Loner-linker LL-Sal4, purified by phenol/chloroform
extraction, and separated from free linkers by
Centricon-100 column. Then, the cDNAs were amplified by
long-range high fidelity PCR using Ex Taq polymerase
(Takara) with a primer Sal4-S for 25 cycles. The products
were purified by phenol/chloroform extraction and
Centricon-100 column. The cDNAs were digested with SalI
and NotI enzymes and cloned into SalI/NotI site of
pCMV-Sport6 plasmid vector. The average insert size is
about 3.6kb."

ORIGIN

Query Match 64.8%; Score 682.8; DB 6; Length 804;
Best Local Similarity 97.3%; Pred. No. 1.5e-170;
Matches 716; Conservative 0; Mismatches 17; Indels 3; Gaps 2;

QY 212 TATCAAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAAAGTACTTTCATCTCTC 271
Db |||||
1 TATCAAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAAAGTACTTTCATCTCTC 60
QY 272 ACATCTTAGC- CTTTTCCTCAGCCAGTGTAAATGTAATGAAATTTGGTGGAGCGC 330
Db |||||
61 ACATCTTAGCNCCTTTTTCAGCCAGTGTAAATGTAATGAAATTTGGTGGAGCGC 120
QY 331 TTTAGTCAGAGGTGCGAGTTCAGAGGCTCGCTGTTCTATGGCTTTCAAAATTCATC 390
Db |||||
121 TTTAGTCAGAGGTGCGAGTTCAGAGGCTCGCTGTTCTATGGCTTTCAAAATTCATC 180
QY 391 GAGAATGTTTCACTCAGAGATGTACAGTTCGTGTATAGACACTTTACATCAGAGATCCCAAG 450
Db |||||
181 GAGAATGTTTCACTCAGAGATGTACAGTTCGTGTATAGACACTTTACATCAGAGATCCCAAG 240
QY 451 AAAAGGAAATTTTATTTAATGCAATGAAACCATGCCCTATGTTAAGAAAAAAGCAGAT 510
Db |||||
241 AAAAGGAAATTTTATTTAATGCAATGAAACCATGCCCTATGTTAAGAAAAAAGCAGAT 300
QY 511 TGGGCCCTCGATGATAGACAGATAGAAAATCTACTTTTGGGGAAGAGTGGTGGCCTTT 570
Db |||||
301 TGGGCCCTCGATGATAGACAGATAGAAAATCTACTTTTGGGGAAGAGTGGTGGCCTTT 360
QY 571 GCTGCTGTAGAGGATTTTCTCTCAGGATCTTTTGTGCTATATTTCTGGCTTAAAGAAG 630
Db |||||
361 GCTGCTGTAGAGGATTTTCTCTCAGGATCTTTTGTGCTATATTTCTGGCTTAAAGAAG 420
QY 631 AGAGGTCTTATGCCAGGACTCACTTTTCCAAATGAACATCATCAGCAGAGATGAAGACTT 690
Db |||||
421 AGAGGTCTTATGCCAGGACTCACTTTTCCAAATGAACATCATCAGCAGAGATGAAGACTT 480
QY 691 CACTGTGACTTTCCTTGCCTGATGTTCCAAATCTTAGTAAATGAAGCCTTCAGAGAAAGG 750
Db |||||
481 CACTGTGACTTTCCTTGCCTGATGTTCCAAATCTTAGTAAATGAAGCCTTCAGAGAAAGG 540
QY 751 GTCAGGAGATCATTTGTCATGCTGTCAAAATGAGCAGAGTCTTTTACAGAGCCTTG 810
Db |||||
541 GTCAGGAGATCATTTGTCATGCTGTCAAAATGAGCAGAGTCTTTTACAGAGCCTTG 600
QY 811 CCAGTTGGCCTCATTTGGAATGAATTCATTTTGAAGAACAGTACATTCAGTTGTAGCT 870
Db |||||
601 CCAGTTGGCCTCATTTGGAATGAATTCATTTTGAAGAACAGTACATTCAGTTGTAGCT 660
QY 871 GACAGATTACTGTGGAACCTTGATTCCTCAAGTTTTCAGCAGAGAAATCCTTTGAT 930
Db |||||
661 GACAGATTACTGTGGAACCTTGAT--CTCAAAGGTTTTCAGCAGAGAAATCCTTTGAT 718
QY 931 TTTATGGAACAATTT 946
Db |||||
719 TATGGAACAATTCCTTT 734

RESULT 14
BM801298
LOCUS
DEFINITION
AGNCOURT_6422860 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5586693
5', mRNA sequence.
BM801298
BM801298.1 GI:19118121
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1061)
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12354 row: p column: 22
High quality sequence stop: 631.
Location/Qualifiers
1..1061
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5586693"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_92"
/note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

FEATURES
source

ORIGIN

Query Match 62.1%; Score 653.8; DB 4; Length 1061;
Best Local Similarity 97.4%; Pred. No. 8.9e-163;
Matches 717; Conservative 0; Mismatches 13; Indels 6; Gaps 5;
QY 1 ATGGGCGACCCGGAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 85 ATGGGCGACCCGGAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 144
QY 61 GACACCAAGAAAGTGAATGAAGTCAATGAAGAGCCACTCTAGAAAGAGTTCTCGC 120
Db 145 GACACCAAGAAAGTGAATGAAGTCAATGAAGAGCCACTCTAGAAAGAGTTCTCGC 204
QY 121 CGTTTTGTTCATCTTTCCAAATCCAGTACCTTGATTTTGGAAAATGTATAACAGGCACAG 180
Db 205 CGTTTTGTTCATCTTTCCAAATCCAGTACCTTGATTTTGGAAAATGTATAACAGGCACAG 264
QY 181 GCTTCCTCTTGGACAGCAGAGAGGTGACTTATCAAAGGATCTCCCTCACTGGAAACAAG 240
Db 265 GCTTCCTCTTGGACAGCAGAGAGGTGACTTATCAAAGGATCTCCCTCACTGGAAACAAG 324
QY 241 CTTAAAGCAGATGAGAAAGTACTTCTCATCTCTCATCTTAGCCCTTTTTCAGCCAGCTGAT 300
Db 325 CTTAAAGCAGATGAGAAAGTACTTCTCATCTCTCATCTTAGCCCTTTTTCAGCCAGCTGAT 384
QY 301 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGGTTCCAGAGGCT 360
Db 385 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGGTTCCAGAGGCT 444
QY 361 CGCTGTTTCTATGGCTTTTCAAAATTTCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG 420
Db 445 CGCTGTTTCTATGGCTTTTCAAAATTTCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG 504
QY 421 CTGATAGACACTTATCATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTTGAA 480
Db 505 CTGATAGACACTTATCATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTTGAA 564
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTTCGATGGATAGCAGATAGAAAA 540
Db 565 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTTCGATGGATAGCAGATAGAAAA 624
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGTCTGTAGAAAGAGTTTCTTCTCAGGA 600
Db 625 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGTCTGTAGAAAGAGTTTCTTCTCAGGA 684
QY 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGCTTTATTCGACAGACTCACTTTTTC 660
Db 685 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGCTC-TATGCCAGGACTCACTTTTTC 743
QY 661 AATGAACACTCAT-CAGCAGAGAT-GAAGGACTTCACCTGTGA--CTTTGCTTGCCTGAT-GT 715

Db 744 CATGGACTCATCCACGAGATGGAGGGACTCCCTGTGAACCTTTGGCTTCGCTCATGCT 803

Qy 716 TCCAACTACTTAGTAAA 731

Db 804 TCCAACTACTNAAGAAA 819

RESULT 15

BUI70979	BUI70979	896 bp	mRNA	linear	EST 04-SEP-2002
LOCUS	AGENCOURT_7940338	NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6143754			
DEFINITION	5'. mRNA sequence.				

ACCESSION BUI70979
VERSION BUI70979.1 GI:22684963
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 896)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrived by: The I.M.A.G.E. Consortium (LBNL)

DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

plate: LLAM13467 row: k column: 19
High quality sequence stop: 619.

```
/organism="Homo sapiens"  
/mol_type="mRNA"  
/db_xref="taxon:9606"  
/clone="IMAGE:6143754"
```

```

/issue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 67"
/notes"Organ: eye; Vector: pCMV-SPORT6; Site: 1. NotI:

```

ORIGIN

Query Match	62.0%	Score	653.2	DB	5	Length	896
Best Local Similarity	98.5%	Pred. No.	1.2e-162				
Matches	670	Conservative	0	Mismatches	8	Indels	2
						Gaps	1

QY 1 ATGGGCGACCGGAAAGCGCGGCGGCTGCATCAGCATGAGAGATCATCTTCA 60

db 55 ATGGGCGACCGGAAAGCGCGGCGGCTGCATCAGCATGAGAGATCATCTTCA 114

61 GACACCAACCAAGTGAAATAAAGTCAAATGAAGGCCACTCCTTAAGAAGAGTTCTCGC 120

db 115 GACACCAACCAAGTGAAATAAAGTCAAATGAAGGCCACTCCTTAAGAAGAGTTCTCGC 174

121 CGGTTTGTCATCTTTCCAATCCAGTACCCCTGATATTTGGAAAATGTATAAACAGGCACAG 180
|||||
175 CGGTTTGTCATCTTTCCAATCCAGTACCCCTGATATTTGGAAAATGTATAAACAGGCACAG 234
|||||

QY	181	GCCTTCCTTCGACAGCAGAGAGGTTGACTTATCAAGGATCTCCCTCACTGGACAAG	240
Db	235	GCCTTCCTTCGACAGCAGAGAGGTCGACTTATCAAGGATCTCCCTCACTGGACAAG	294

241	CTTTAAGCAGATGAGAGTACTTCATCTCTCATCTTAGCCCTTTTTCAGCCAGTGAT	300
QY		
db	295 CTTTAAAGCAGATGAGAGTACTTCATCTCTCATCTTAGCCCTTTTTCAGCCAGTGAT	354

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Qy		301	GGAATTGTAATGAAATTTGGTGGAGCGCTTTTAGTCAGGAGGTCCAGAGGCT	360
Db		355	GGAATTGTAATGAAATTTGGTGGAGCGCTTTTAGTCAGGAGGTCCAGAGGCT	414
Qy		361	CGCTGTTTTCTATGCGTTTCAAATCTCATCGAGATGTTCACTCAGAGNTGTACAGTTTG	420
Db		415	CGCTGTTTTCTATGCGTTTCAAATCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG	474
Qy		421	CTGATAGACACTTACATCAGAGATCCCAGAAGAAAGGGAATTTTATTTAATGCAATTGAA	480
Db		475	CTGATAGACACTTACATCAGAGATCCCAGAAGAAAGGGAATTTTATTTAATGCAATTGAA	534
Qy		481	ACCATGCCCTATTGTTAAGAAAAAGACAGATTGGGCCTTGGCATGGATAGCAGATAGAAA	540
Db		535	ACCATGCCCTATTGTTAAGAAAAAGACAGATTGGGCCTTGGCATGGATAGCAGATAGAAA	594
Qy		541	TCTACTTTTTGGGAAACAGTGGTGGCCCTTCTGCTGTAGAGAGAGTTTCTTCTCAGGA	600
Db		595	TCTACTTTTTGGGAAACAGTGGTGGCCCTTCTGCTGTAGAGAGAGTTTCTTCTCAGGA	654
Qy		601	TCTTTTCTGCTATTCTTGCTTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTT--	658
Db		655	TC TT T T G C T A T T C T G C T A A G A A G A G A G G T C T A T G C C C A G A C T C A C T T T T T C	714
Qy		659	CCAATGAAC TC AT CAG CAGA 678	
Db		715	CAATGAAACT CAT CAG CAGA 734	

Search completed: October 30, 2005, 03:23:17
Job time : 2721.5 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	1053	100.0	1053	6	AR454868	Sequence
2	1053	100.0	1053	6	BD064764	Novel pro
3	1053	100.0	1053	6	BD093077	Novel pro
4	1053	100.0	1056	6	CQ714252	Sequence
5	1053	100.0	1056	9	AB166671	Homo sapi
6	1053	100.0	1055	6	AR454869	Sequence
7	1053	100.0	4955	6	BD064765	Novel pro
8	1053	100.0	4955	6	BD093078	Novel pro
9	1053	100.0	4955	9	AB036063	Homo sapi
10	1051.4	99.8	1053	6	AR454877	Sequence
11	1051.4	99.8	1053	6	BD064773	Novel pro
12	1051.4	99.8	1053	6	BD093086	Novel pro
13	1051.4	99.8	1081	6	AR454870	Sequence
14	1051.4	99.8	1081	6	BD064766	Novel pro
15	1051.4	99.8	1081	6	BD093079	Novel pro
16	1049.8	99.7	1601	6	BD156916	Primer fo
17	1049.8	99.7	1601	6	AR877905	Sequence
18	1049.8	99.7	1601	9	AK001965	Homo sapi
19	895.2	85.0	4532	10	BC058103	Mus muscu

```
Db 241 CTTAAAGCAGATGAGAGTACTTCTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTGAT 300
Qy 301 GGAATTTGAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTCAGGTTCCAGAGCT 360
Db 301 GGAATTTGAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTCAGGTTCCAGAGCT 360
Qy 361 CGCTGTTTCTATGCTTTCAAATTTCTCATCAGAAATGTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGCTTTCAAATTTCTCATCAGAAATGTTCACTCAGAGATGTACAGTTTG 420
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTAAATGCAATTTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTAAATGCAATTTGAA 480
Qy 481 ACCATGCCCTATGTTAAGAAAAGCAGATGGGCGCTTGCATGATGATGAGATAGAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAAGCAGATGGGCGCTTGCATGATGATGAGATAGAAA 540
Qy 541 TCTACTTTTGGGAAAAGAGTGGTGGCGCTTCTGCTGTAGAGGAGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAAGAGTGGTGGCGCTTCTGCTGTAGAGGAGTTTCTTCTCAGGA 600
Qy 601 TCTTTTGTGCTATATTTCTGCTTAAAGAGAGAGTCTTATGCCAGGACTCCTTTTTC 660
Db 601 TCTTTTGTGCTATATTTCTGCTTAAAGAGAGAGTCTTATGCCAGGACTCCTTTTTC 660
Qy 661 AATGAATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA 720
Db 661 AATGAATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA 720
Qy 721 TACTTAGTAAATAGCCTTCAAGAAAGGTCAGGAGATCATTTGTGCTGTCAA 780
Db 721 TACTTAGTAAATAGCCTTCAAGAAAGGTCAGGAGATCATTTGTGCTGTCAA 780
Qy 781 ATTGAGCAGGAGTTTAAACAGAGCCCTTGCCAGTGGCCCTCATTTGGAATGAAATTGCA 840
Db 781 ATTGAGCAGGAGTTTAAACAGAGCCCTTGCCAGTGGCCCTCATTTGGAATGAAATTGCA 840
Qy 841 TTGATGAAACAGTACATTTGTTGTAGTCAGCAGATTTACTTTGGAACTTTGGATTTCA 900
Db 841 TTGATGAAACAGTACATTTGTTGTAGTCAGCAGATTTACTTTGGAACTTTGGATTTCA 900
Qy 901 AAGGTTTTTCAGGAGAGAAATCCCTTTGATTTTATGGAACATTTCTTTAGAGGAAA 960
Db 901 AAGGTTTTTCAGGAGAGAAATCCCTTTGATTTTATGGAACATTTCTTTAGAGGAAA 960
Qy 961 ACAAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTCAGGATTTAGGCAAAACC 1020
Db 961 ACAAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTCAGGATTTAGGCAAAACC 1020
Qy 1021 ACAGATAACGCTTTTCACTTGGATGCGAGATTTT 1053
Db 1021 ACAGATAACGCTTTTCACTTGGATGCGAGATTTT 1053
```

RESULT 2

```
BD064764
LOCUS BD064764
DEFINITION Novel protein and DNA thereof.
ACCESSION BD064764
VERSION BD064764.1 GI:22610367
KEYWORDS JP 2001269184-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1053)
Nakamura,Y., Arakawa,H. and Tanaka,H.
Novel protein and DNA thereof
Patent: JP 2001269184-A 1 02-Oct-2001;
YUSUKE NAKAMURA, TAKEDA CHEMICAL INDUSTRIES LTD
OS Homo sapiens (human)
```

```
PN JP 2001269184-A/1
PD 02-OCT-2001
PF 27-JUN-2000 JP 2000192401
PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
C12N15/09, A61K31/711, A61K38/00, A61K38/43, A61K45/00, A61K48/00, PC
A61P35/00,
PC A61P43/00, C07K16/32, C07K16/40, C12N1/15, C12N1/19, C12N1/21 PC
C12N5/10, C12N9/02,
PC C12P21/02, C12P21/08, G01N33/15, G01N33/50, G01N33/53// (C12N15/09,
C12R1/91),
PC (C12N15/09, C12R1/19), (C12N9/02, C12R1/91), (C12P21/02, C12R1/91),
PC (C12P21/08, C12R1/91), (C12N15/00, A61K37/02, A61K37/48, C12N5/00,
C12N15/00, C12R1/91), (C12N15/00, C12R1/19)
CC Novel protein and DNA thereof
FH Key Location/Qualifiers
FT source 1..1053
FT /organism="Homo sapiens (human)".
FEATURES
source
1..1053
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 1053; DB 6; Length 1053;
Best Local Similarity 100.0%; Pred. No. 1.2e-255;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGGCGACCGGAAAGCGCGGAGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 ATGGGCGACCGGAAAGCGCGGAGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Qy 61 GACACCAACCAAGCTGAAATTAAGTCAAAATGAAGGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAACCAAGCTGAAATTAAGTCAAAATGAAGGCCACTCTTAAGAAAGAGTTCTCGC 120
Qy 121 CGGTTTGTCTCTTTCCAAATCCAGTACCTCATATTTGGAATAATGTAATAACAGGCACAG 180
Db 121 CGGTTTGTCTCTTTCCAAATCCAGTACCTCATATTTGGAATAATGTAATAACAGGCACAG 180
Qy 181 GCTTCTCTTGGACAGCAGAGAGGTGACATTAACAAGGATCTCCCTCACTGGAAACAAG 240
Db 181 GCTTCTCTTGGACAGCAGAGAGGTGACATTAACAAGGATCTCCCTCACTGGAAACAAG 240
Qy 241 CTTAAACAGATGAGAGTACTTCTATCTCAGATCTTAGCCTTTTTCAGCCAGTGTAT 300
Db 241 CTTAAACAGATGAGAGTACTTCTATCTCAGATCTTAGCCTTTTTCAGCCAGTGTAT 300
Qy 301 GGAATTTGAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTCAGGTTCCAGAGCT 360
Db 301 GGAATTTGAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTCAGGTTCCAGAGCT 360
Qy 361 CGCTGTTTCTATGCTTTCAAATTTCTCATCAGAAATGTTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGCTTTCAAATTTCTCATCAGAAATGTTTCACTCAGAGATGTACAGTTTG 420
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTAAATGCAATTTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTAAATGCAATTTGAA 480
Qy 481 ACCATGCCCTATGTTAAGAAAAGCAGATGGGCGCTTGCATGATGATGAGATAGAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAAGCAGATGGGCGCTTGCATGATGATGAGATAGAAA 540
Qy 541 TCTACTTTTGGGAAAAGAGTGGTGGCGCTTCTGCTGTAGAGGAGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAAGAGTGGTGGCGCTTCTGCTGTAGAGGAGTTTCTTCTCAGGA 600
Qy 601 TCTTTTGTGCTATATTTCTGCTTAAAGAGAGTCTTATGCCAGGACTCCTTTTTC 660
Db 601 TCTTTTGTGCTATATTTCTGCTTAAAGAGAGTCTTATGCCAGGACTCCTTTTTC 660
Qy 661 AATGAATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA 720
Db 661 AATGAATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA 720
Qy 721 TACTTAGTAAATAGCCTTCAAGAAAGGTCAGGAGATCATTTGTGCTGTCAA 780
Db 721 TACTTAGTAAATAGCCTTCAAGAAAGGTCAGGAGATCATTTGTGCTGTCAA 780
Qy 781 ATTGAGCAGGAGTTTAAACAGAGCCCTTGCCAGTGGCCCTCATTTGGAATGAAATTGCA 840
Db 781 ATTGAGCAGGAGTTTAAACAGAGCCCTTGCCAGTGGCCCTCATTTGGAATGAAATTGCA 840
Qy 841 TTGATGAAACAGTACATTTGTTGTAGTCAGCAGATTTACTTTGGAACTTTGGATTTCA 900
Db 841 TTGATGAAACAGTACATTTGTTGTAGTCAGCAGATTTACTTTGGAACTTTGGATTTCA 900
Qy 901 AAGGTTTTTCAGGAGAGAAATCCCTTTGATTTTATGGAACATTTCTTTAGAGGAAA 960
Db 901 AAGGTTTTTCAGGAGAGAAATCCCTTTGATTTTATGGAACATTTCTTTAGAGGAAA 960
Qy 961 ACAAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTCAGGATTTAGGCAAAACC 1020
Db 961 ACAAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTCAGGATTTAGGCAAAACC 1020
Qy 1021 ACAGATAACGCTTTTCACTTGGATGCGAGATTTT 1053
Db 1021 ACAGATAACGCTTTTCACTTGGATGCGAGATTTT 1053
```

```

Db      661  AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCTGCTGATGTTCCAA 720
Qy      721  TACTTAGTAAATAAGCCTTTCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCTCAA 780
Db      721  TACTTAGTAAATAAGCCTTTCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCTCAA 780
Qy      781  ATTGAGCAGGAGTTTAAACAGAGCCCTTGCCAGTTCAGTTCGCTCATTTGGAAATGAATTCGATT 840
Db      781  ATTGAGCAGGAGTTTAAACAGAGCCCTTGCCAGTTCAGTTCGCTCATTTGGAAATGAATTCGATT 840
Qy      841  TTGATGAACAGTACATTTGTTGATGCTCAGATTCAGATTCAGTTCGCTCATTTGGAAATGAATTCGATT 900
Db      841  TTGATGAACAGTACATTTGTTGATGCTCAGATTCAGATTCAGTTCGCTCATTTGGAAATGAATTCGATT 900
Qy      901  AAGGTTTTTTCAGGAGAGAAATCCCTTTGATTTTATGAAACATTTCTTTTGAAGAGGAAAA 960
Db      901  AAGGTTTTTTCAGGAGAGAAATCCCTTTGATTTTATGAAACATTTCTTTTGAAGAGGAAAA 960
Qy      961  ACAAAATTTCTTTGAGAAACAGATTTTCAGATTCAGGCTTTGCGATTCAGGCTTTGCGATTCGCGA 1020
Db      961  ACAAAATTTCTTTGAGAAACAGATTTTCAGATTCAGGCTTTGCGATTCAGGCTTTGCGATTCGCGA 1020
Qy      1021  ACAGATAACGCTTTCACCTTGCGATTCGAGATTTT 1053
Db      1021  ACAGATAACGCTTTCACCTTGCGATTCGAGATTTT 1053

```

RESULT 3

```

BD093077
LOCUS      BD093077      1053 bp      DNA      linear      PAT 27-AUG-2002
DEFINITION Novel protein and its DNA.
ACCESSION BD093077
VERSION    BD093077.1 GI:22638665
KEYWORDS  WO 0100799-A/1.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 1053)
AUTHORS    Nakamura,Y., Arakawa,H. and Tanaka,H.
TITLE      Novel protein and its DNA
JOURNAL    Patent: WO 0100799-A 1 04-JAN-2001;
            TAKEDA CHEMICAL INDUSTRIES LTD, YUSUKE NAKAMURA, HIROFUMI ARAKAWA,
            HIROSHI TANAKA
COMMENT     OS Homo sapiens (human)
            PN WO 0100799-A/1
            PD 04-JAN-2001
            PF 27-JUN-2000 WO 2000JP004189
            PR 28-JUN-1999 JP 99P 181131,06-JUL-1999 JP 99P 192391 PR
            21-JAN-2000 JP 00P 017770
            PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
            C12N9/04, C12N15/53, C12N1/15, C12N1/19, C12N5/10 PC
            ,C12P21/02, A61K38/44,
            PC A61K48/00, C07K16/40, G01N33/53, C12Q1/26
            CC
            FH

```

FEATURES

```

source      Location/Qualifiers
            1..1053
               /organism="Homo sapiens"
               /mol_type="genomic DNA"
               /db_xref="taxon:9606"

```

ORIGIN

```

Query Match      100.0%; Score 1053; DB 6; Length 1053;
Beet Local Similarity 100.0%; Pred. No. 1.2e-255;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  ATGGCGACCGGAAAGCCGGAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db      1  ATGGCGACCGGAAAGCCGGAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Qy      61  GACACCAAGAACTGAATAAGTCAATGAAGAGCCACTCTCAAGAAAGAGTTCTCGC 120

```

RESULT 4
 CQ714252
 LOCUS
 DEFINITION Sequence 186 from Patent WO02068579.
 ACCESSION CQ714252

CQ714252 1056 bp DNA linear PAT 03-FEB-2004
 Sequence 186 from Patent WO02068579.
 CQ714252

```

Db      61  GACACCAAGAAAGTGAATAAGTCAATAAGAGGCCACTCTTAAGAAAGAGTTCTCGC 120
Qy      121  CGGTTTGTTCATCTTTCCAAATCCAGTACCTCATATTTGGAAAAATGTATAAA CAGGCACAG 180
Db      121  CGGTTTGTTCATCTTTCCAAATCCAGTACCTCATATTTGGAAAAATGTATAAA CAGGCACAG 180
Qy      181  GCTTCTCTTCTGGACAGCAGAGAGGTGAGATTTATCAAGAGATCTCCCTCCTCTGGAACAG 240
Db      181  GCTTCTCTTCTGGACAGCAGAGAGGTGAGATTTATCAAGAGATCTCCCTCCTCTGGAACAG 240
Qy      241  CTTAAAGCAGATGAGAGTACTTCTCATCTCTCAATCTTTAGGCTTTTTCAGGCGCAGTGAT 300
Db      241  CTTAAAGCAGATGAGAGTACTTCTCATCTCTCAATCTTTAGGCTTTTTCAGGCGCAGTGAT 300
Qy      301  GGAATGCTTAATGAATAATTTGGTGGAGCGCTTTAGTCAAGGAGGTGAGGTTCCAGAGCT 360
Db      301  GGAATGCTTAATGAATAATTTGGTGGAGCGCTTTAGTCAAGGAGGTGAGGTTCCAGAGCT 360
Qy      361  CGCTGTTTCTATGCTTTCAAAATTTCTCATCAGAGATGTTCACTCAGAGATGTACAGTTTG 420
Db      361  CGCTGTTTCTATGCTTTCAAAATTTCTCATCAGAGATGTTCACTCAGAGATGTACAGTTTG 420
Qy      421  CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAATGCAATTGAA 480
Db      421  CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAATGCAATTGAA 480
Qy      481  ACCATGCGCTATGTTAAGAAAAAGCAGATTTGGCGCTTGCATGATGAGATGAGAAAA 540
Db      481  ACCATGCGCTATGTTAAGAAAAAGCAGATTTGGCGCTTGCATGATGAGATGAGAAAA 540
Qy      541  TCTACTTTTGGGAAAGAGTGTGCTGCTGTGTAGAGAGGATTTTCTTCTCAGCA 600
Db      541  TCTACTTTTGGGAAAGAGTGTGCTGCTGTGTAGAGAGGATTTTCTTCTCAGCA 600
Qy      601  TCTTTTGTCTGCTATATTCTGGCTTAAAGAGAGAGTCTTATGCCAGGACTCCTTTTTC 660
Db      601  TCTTTTGTCTGCTATATTCTGGCTTAAAGAGAGAGTCTTATGCCAGGACTCCTTTTTC 660
Qy      661  AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTGCTGATGTTCCAA 720
Db      661  AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTGCTGATGTTCCAA 720
Qy      721  TACTTAGTAAATAAGCCTTTCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 780
Db      721  TACTTAGTAAATAAGCCTTTCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 780
Qy      781  ATTGAGCAGGAGTTTAAACAGAGCCCTTGCCAGTTCGCTCATTTGGAATGAATTCGATT 840
Db      781  ATTGAGCAGGAGTTTAAACAGAGCCCTTGCCAGTTCGCTCATTTGGAATGAATTCGATT 840
Qy      841  TTGATGAACAGTACATTTGATGTTTGTAGTGCACAGATTTACTTTGGAACTTTGGATTCTCA 900
Db      841  TTGATGAACAGTACATTTGATGTTTGTAGTGCACAGATTTACTTTGGAACTTTGGATTCTCA 900
Qy      901  AAGTTTTTTCAGGAGAGAAATCCCTTTGATTTTATGAAACATTTCTTTTGAAGGAAAA 960
Db      901  AAGTTTTTTCAGGAGAGAAATCCCTTTGATTTTATGAAACATTTCTTTTGAAGGAAAA 960
Qy      961  ACAAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGGCTTTGCGATTCAGGCTTTGCGA 1020
Db      961  ACAAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGGCTTTGCGATTCAGGCTTTGCGA 1020
Qy      1021  ACAGATAACGCTTTCACCTTGCGATTCGAGATTTT 1053
Db      1021  ACAGATAACGCTTTCACCTTGCGATTCGAGATTTT 1053

```

```

VERSION      CQ714252.1  GI:42275109
KEYWORDS
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1
AUTHORS      Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE        Kits, such as nucleic acid arrays, comprising a majority of
             humanexons or transcripts, for detecting expression and other uses
             thereof
JOURNAL      Patent: WO 02068579-A 186 06-SEP-2002;
             PE Corporation (NY) (US)
FEATURES     Location/Qualifiers
             source
               1..1056
               /organism="Homo sapiens"
               /mol_type="unassigned DNA"
               /db_xref="taxon:9606"
ORIGIN
Query Match      100.0%; Score 1053; DB 6; Length 1056;
Best Local Similarity 100.0%; Pred. No. 1.2e-255;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGGCGACCGGAAAGCGCGGAGCGGCTGGATCAGATGAGATGATCATCTTCA 60
Db 1 ATGGGCGACCGGAAAGCGCGGAGCGGCTGGATCAGATGAGATGATCATCTTCA 60
Qy 61 GACACCAACGAACTGAATTAAGTCAATTAAGAGCCACTCCTTAAGAAAGATTCGCG 120
Db 61 GACACCAACGAACTGAATTAAGTCAATTAAGAGCCACTCCTTAAGAAAGATTCGCG 120
Qy 121 CGGTTGTCATCTTCCAAATCCAGTACCTGATATTTGGAAATGTATAACAGGCACAG 180
Db 121 CGGTTGTCATCTTCCAAATCCAGTACCTGATATTTGGAAATGTATAACAGGCACAG 180
Qy 181 GCTTCTCTTGACAGCAGAGAGGTGCACTTATCAAGAGGATCTCCCTCACTGGAACAAG 240
Db 181 GCTTCTCTTGACAGCAGAGAGGTGCACTTATCAAGAGGATCTCCCTCACTGGAACAAG 240
Qy 241 CTTAAACAGATGAGATGACTTCACTCTCACATCTTAGCTTTTGGACGAGTGAT 300
Db 241 CTTAAACAGATGAGATGACTTCACTCTCACATCTTAGCTTTTGGACGAGTGAT 300
Qy 301 GGAATCTGAATGAAATTTGGTGAGCGCTTTAGTCAGAGGTCGAGGTTCCAGAGGCT 360
Db 301 GGAATCTGAATGAAATTTGGTGAGCGCTTTAGTCAGAGGTCGAGGTTCCAGAGGCT 360
Qy 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGATGTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGATGTTCACTCAGAGATGTACAGTTTG 420
Qy 421 CTGATAGACATTAATCATCAGATGCCAAGAAAGGGAATTTTATTTAATGCAATTTGAA 480
Db 421 CTGATAGACATTAATCATCAGATGCCAAGAAAGGGAATTTTATTTAATGCAATTTGAA 480
Qy 481 ACCATGCCCTATGTTAAGAAAGACAGATTTGGCCCTTGGATGATGATGATAGAGAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAGACAGATTTGGCCCTTGGATGATGATGATAGAGAAA 540
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCCCTTTGCTCTGTAGAGAGGATTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCCCTTTGCTCTGTAGAGAGGATTTTCTTCTCAGGA 600
Qy 601 TCTTTTCTGCTATATTTCTGCTAAGAGAGAGGTCCTTATGCCAGGACTCACTTTTCC 660
Db 601 TCTTTTCTGCTATATTTCTGCTAAGAGAGAGGTCCTTATGCCAGGACTCACTTTTCC 660
Qy 661 AATGAATCATCAGCAGATGAGGATTCACATGTCATTTGCTTGCCTGATGTTCCAA 720
Db 661 AATGAATCATCAGCAGATGAGGATTCACATGTCATTTGCTTGCCTGATGTTCCAA 720
Qy 721 TACTTAGTAAATAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGTTGATGTCGTCAA 780
Db 721 TACTTAGTAAATAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGTTGATGTCGTCAA 780

Db 721 TACTTAGTAAATAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGTTGATGTCGTCAA 780
Qy 781 ATTGAGCAGGAGTATTTTAAACAGAAAGCCCTTGCAGTTGGCCCTCATTTGGAATGATTTGCATT 840
Db 781 ATTGAGCAGGAGTATTTTAAACAGAAAGCCCTTGCAGTTGGCCCTCATTTGGAATGATTTGCATT 840
Qy 841 TTGATGAAACAGTACATTTGAGTTTGTAGCTGACAGATTTACTTTGTGGAACCTTGGATTCTCA 900
Db 841 TTGATGAAACAGTACATTTGAGTTTGTAGCTGACAGATTTACTTTGTGGAACCTTGGATTCTCA 900
Qy 901 AAGGTTTTTTCAGGCAGAAATCCTTTTGAATTTTATGGAATAACATTTCTTTAGAGAGAAA 960
Db 901 AAGGTTTTTTCAGGCAGAAATCCTTTTGAATTTTATGGAATAACATTTCTTTAGAGAGAAA 960
Qy 961 ACAAATTTCTTTGAGAAACGAGTTTTCAGATGATCAGCGTTTTCAGAGTTATGTCAGAGAAC 1020
Db 961 ACAAATTTCTTTGAGAAACGAGTTTTCAGATGATCAGCGTTTTCAGAGTTATGTCAGAGAAC 1020
Qy 1021 ACAGATAACGTTCTTCACTTGCATGCGAGATTTT 1053
Db 1021 ACAGATAACGTTCTTCACTTGCATGCGAGATTTT 1053

RESULT 5
AB166671
LOCUS       AB166671
DEFINITION Homo sapiens p53R2 mRNA for p53-inducible ribonucleotide reductase
             small subunit 2, complete cds.
ACCESSION   AB166671
VERSION      AB166671.1  GI:45259568
KEYWORDS     Homo sapiens (human)
SOURCE       Homo sapiens
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1
AUTHORS      Ugai, H. and Yokoyama, K.K.
TITLE        Homo sapiens p53R2 mRNA for p53-inducible ribonucleotide reductase
             small subunit 2, complete cds
JOURNAL      Published Only in Database (2004)
REFERENCE    2 (bases 1 to 1056)
AUTHORS      Ugai, H. and Yokoyama, K.K.
TITLE        Direct Submission
JOURNAL      Submitted (05-MAR-2004) Hideyo Ugai, RIKEN BioResource Center, Gene
             Engineering Division; 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074,
             Japan (E-mail:ugai@rtc.riken.go.jp, Tel:81-29-836-3612,
             Fax:81-29-836-9120)
FEATURES     Location/Qualifiers
             1..1056
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="p53R2.C19"
             /cell_line="Saos-2"
             /cell_type="Osteogenic sarcoma"
             1..1056
             /gene="p53R2"
             1..1056
             /gene="p53R2"
             /codon_start=1
             /product="p53-inducible ribonucleotide reductase small
             subunit 2"
             /protein_id="BAD12267.1"
             /db_xref="GI:45259569"
             /translation="MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLRKSSRFVI
             PFIQYDPIWMYKQAOASFABEVDLSKDLPHWNKLKADSKYFTSHLAPFAASDGI
             VNENLVERFQEVQVPEARCFYQIILNIENVHSEMSYSLIDTYIDDPKRELFNAIE
             TMPYKKADWALRWIADRAKSTGERVVAFAAVGVPFSGFAALFWLKKGLMFGLT
             FSNELISRDEGLCHDFACLMFQYLVNKPSESRVREIIVDAVKIIOEFLEALPVLGIG
             MNCILMKQYIEFVADRLILBELGFSKQFQAEINPFDMENISLEKTNFFEKRVSEYQF
             AVMAETTDNVTFLDADF"
ORIGIN

```

Query Match	100.0%	Score 1053;	DB 9;	Length 1056;
Best Local Similarity	100.0%;	Prod. No. 1.2e-255;		
Matches 1053;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	1	ATGGGCGACCGGAAGCGCGGAAGCGGCGGGCTGGATCAGGATGAGAGATCACTTCA	60
Db	1	ATGGGCGACCGGAAGCGCGGAAGCGGCGGGCTGGATCAGGATGAGAGATCACTTCA	60
Qy	61	GACACCAACGAAGTGAAATATAAGTCAATGGAAGGCCACTCTCTAAGAAAGAGTTCTCGC	120
Db	61	GACACCAACGAAGTGAAATATAAGTCAATGGAAGGCCACTCTCTAAGAAAGAGTTCTCGC	120
Qy	121	CGGTTTGTCTATCTTTTCCAAATCCAGTACCCCTGATATTTGGAAAATGTATAAACAGGCAAG	180
Db	121	CGGTTTGTCTATCTTTTCCAAATCCAGTACCCCTGATATTTGGAAAATGTATAAACAGGCAAG	180
Qy	181	GCTTCTCTTGTGACAGCAGAGAAGGTCGACTTATCAAAAGGATCTCCCTCACTGGAACAAG	240
Db	181	GCTTCTCTTGTGACAGCAGAGAAGGTCGACTTATCAAAAGGATCTCCCTCACTGGAACAAG	240
Qy	241	CTTAAAGCAGATGAGAAGTACTTCACTCTCACATCTTAGCCTTTTTTCGAGCCAGTGAT	300
Db	241	CTTAAAGCAGATGAGAAGTACTTCACTCTCACATCTTAGCCTTTTTTCGAGCCAGTGAT	300
Qy	301	GGAAATGTAAATGAAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT	360
Db	301	GGAAATGTAAATGAAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT	360
Qy	361	CGCTGTTTCTATGGCTTTTCAAATCTCATCGAAGATGTTCACTCAGAGATGTAACAGTTTG	420
Db	361	CGCTGTTTCTATGGCTTTTCAAATCTCATCGAAGATGTTCACTCAGAGATGTAACAGTTTG	420
Qy	421	CTGATGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTTTATTAATGCAATTTGAA	480
Db	421	CTGATGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTTTATTAATGCAATTTGAA	480
Qy	481	ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGGCGCTTGCATGGATAGCAGATAGAAAA	540
Db	481	ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGGCGCTTGCATGGATAGCAGATAGAAAA	540
Qy	541	TCTACTTTTGGGAAAGAGTGGTGGCGCTTGTCTGTAGAGGAGTTTTCTTCTCAGGA	600
Db	541	TCTACTTTTGGGAAAGAGTGGTGGCGCTTGTCTGTAGAGGAGTTTTCTTCTCAGGA	600
Qy	601	TCTTTTGTCTGTATATTCGGCTTAAAGAGAGAGGCTTATGCGAGACTCACTTTTTC	660
Db	601	TCTTTTGTCTGTATATTCGGCTTAAAGAGAGAGGCTTATGCGAGACTCACTTTTTC	660
Qy	661	AATGAACTCATCAGCAGAGATGAAGACTTTCACCTGTGACCTTTCCTGCGCTGATGTTCCAA	720
Db	661	AATGAACTCATCAGCAGAGATGAAGACTTTCACCTGTGACCTTTCCTGCGCTGATGTTCCAA	720
Qy	721	TACTTAGTAAATAAAGCCCTTCAGAGAAGAGGGTTCAGGAGATCAATGTTGATGCTGCAAA	780
Db	721	TACTTAGTAAATAAAGCCCTTCAGAGAAGAGGGTTCAGGAGATCAATGTTGATGCTGCAAA	780
Qy	781	ATTGAGCAGAGATTTTTTAAACAGAACCTTCGCCATGTTGGCCCTCAATGGAATGAATTCGATT	840
Db	781	ATTGAGCAGAGATTTTTTAAACAGAACCTTCGCCATGTTGGCCCTCAATGGAATGAATTCGATT	840
Qy	841	TTGATGAAAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTCTGGAACCTTGATTTCTCA	900
Db	841	TTGATGAAAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTCTGGAACCTTGATTTCTCA	900
Qy	901	AAGGTTTTTTCAGGCAGAAAAATCCCTTTTGATTTTTATGAAAAACATTTCTTTAGAGGAAAA	960
Db	901	AAGGTTTTTTCAGGCAGAAAAATCCCTTTTGATTTTTATGAAAAACATTTCTTTAGAGGAAAA	960
Qy	961	ACAAATTTCTTTGAGAAAACGAGTTTCAGAGTATCAGCGTTTTTCGAGTTATGCGAGAAACC	1020
Db	961	ACAAATTTCTTTGAGAAAACGAGTTTCAGAGTATCAGCGTTTTTCGAGTTATGCGAGAAACC	1020

Qy	1021	ACGATAAACGCTTTCACCTTGGATGCAGATTTT	1053
Db	1021	ACAGATAACGCTTTCACCTTGGATGCAGATTTT	1053
RESULT 6			
AR454869			
LOCUS	AR454869	4955 bp	linear
DEFINITION	Sequence 3 from patent US 6682917.		
ACCESSION	AR454869		
VERSION	AR454869.1	GI:42688824	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 4955)		
AUTHORS	Nakamura, Y., Arakawa, H. and Tanaka, H.		
TITLE	Protein having a ribonucleotide Reductase activity and a DNA thereof		
JOURNAL	Patent: US 6682917-A 3 27-JAN-2004;		
FEATURES	Location/Qualifiers		
source	1..4955		
	/organism="unknown"		
	/mol_type="genomic DNA"		
ORIGIN			
	Query Match	100.0%;	Score 1053; DB 6; Length 4955;
	Best Local Similarity	100.0%;	Pred. No. 1.1e-255;
	Matches 1053; Conservative	0; Mismatches	0; Indels 0; Gaps 0;
Qy	1	ATGGGGACCGCGAAAGCGCGAAGCGCGCGCTGGATCAGGATGAGAGATCATCTTCA	60
Db	245	ATGGCGACCGCGAAAGCGCGAAGCGCGCGCTGGATCAGGATGAGAGATCATCTTCA	304
Qy	61	GACACCAAGAAAGTGAATAAAGTCAATGAAGACCCACTCTCTAAGAAAGAGTCTCGC	120
Db	305	GACACCAAGAAAGTGAATAAAGTCAATGAAGACCCACTCTCTAAGAAAGAGTCTCGC	364
Qy	121	CGGTGTGTCACTCTTCCAAATCCAGTACCTGTATTTGGAAATGTATAACAGGCACAG	180
Db	365	CGGTGTGTCACTCTTCCAAATCCAGTACCTGTATTTGGAAATGTATAACAGGCACAG	424
Qy	181	GCCTTCCTTCTGGACACGAGAAGGTGCAGCTTATCAAAGGATCTCCCTCACTGGAAACAAG	240
Db	425	GCCTTCCTTCTGGACACGAGAAGGTGCAGCTTATCAAAGGATCTCCCTCACTGGAAACAAG	484
Qy	241	CTTAAAGCAGATGAGAAGTACTTTCATCTCTCAACCTTAGCCCTTTTTCGAGCCAGTGAT	300
Db	485	CTTAAAGCAGATGAGAAGTACTTTCATCTCTCAACCTTAGCCCTTTTTCGAGCCAGTGAT	544
Qy	301	GGAAATGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTCAGAGGCT	360
Db	545	GGAAATGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTCAGAGGCT	604
Qy	361	CGCTGTTTCTATGGCTTTTCAAAATTTCTCATCGAAGATGTTTCACTCAGAGATGTACAGTTTG	420
Db	605	CGCTGTTTCTATGGCTTTTCAAAATTTCTCATCGAAGATGTTTCACTCAGAGATGTACAGTTTG	664
Qy	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTTTAAATGCAATTGAA	480
Db	665	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTTTAAATGCAATTGAA	724
Qy	481	ACCATGCCCTATGTTAAGAAAAGCAGATTCGGCTTCGGATGGATAGCAGATAGAAA	540
Db	725	ACCATGCCCTATGTTAAGAAAAGCAGATTCGGCTTCGGATGGATAGCAGATAGAAA	784
Qy	541	TCTACTTTTGGGAAAGAGTGTGGCTTTTGTCTGTGTAGAAGGAGTTTTTCTTCAGGA	600
Db	785	TCTACTTTTGGGAAAGAGTGTGGCTTTTGTCTGTGTAGAAGGAGTTTTTCTTCAGGA	844
Qy	601	TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGTCTTTATGCCAGGATCTCACTTTTCC	660
Db	845	TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGTCTTTATGCCAGGATCTCACTTTTCC	904

Qy 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGTGCTGATGTTCCAA 720
Db 905 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGTGCTGATGTTCCAA 964
Qy 721 TACTTAGTAATAAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGTGATGCTGTCAAA 780
Db 965 TACTTAGTAATAAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGTGATGCTGTCAAA 1024
Qy 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCCAGTTCGACGTTTGAATGAATTTGCATT 840
Db 1025 ATTGAGCAGGAGTTTAAACAGAGCCTTGCCAGTTCGACGTTTGAATGAATTTGCATT 1084
Qy 841 TTGATGAACAGTACATTGAGTTTGTAGTCGACAGATTACTTGTGGAACCTTGGATTCTCA 900
Db 1085 TTGATGAACAGTACATTGAGTTTGTAGTCGACAGATTACTTGTGGAACCTTGGATTCTCA 1144
Qy 901 AAGGTTTTTCAGGAGAGAAATCCCTTTGATTTTATGGAAGAAATTTCTTTAGAGAGAAA 960
Db 1145 AAGGTTTTTCAGGAGAGAAATCCCTTTGATTTTATGGAAGAAATTTCTTTAGAGAGAAA 1204
Qy 961 ACAAATTTCTTTGAGAAACAGAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGCAGAAACC 1020
Db 1205 ACAAATTTCTTTGAGAAACAGAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGCAGAAACC 1264
Qy 1021 ACAGATAACGTCCTTTCACCTTGGATGCGAGATTTT 1053
Db 1265 ACAGATAACGTCCTTTCACCTTGGATGCGAGATTTT 1297

RESULT 7
BD064765
LOCUS BD064765 4955 bp DNA linear PAT 27-AUG-2002
DEFINITION Novel protein and DNA thereof.
ACCESSION BD064765
VERSION BD064765.1 GI:22610368
KEYWORDS JP 2001269184-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4955)
AUTHORS Nakamura,Y., Arakawa,H. and Tanaka,H.
TITLE Novel protein and DNA thereof
JOURNAL Patent: JP 2001269184-A 2 02-OCT-2001;
YUSUKE NAKAMURA,TAKEDA CHEMICAL INDUSTRIES LTD
COMMENT OS Homo sapiens (human)
PN JP 2001269184-A/2
PD 02-OCT-2001
PF 27-JUN-2000 JP 2000192401
PI YUSUKE NAKAMURA,HIROFUMI ARAKAWA,HIROSHI TANAKA PC
C12N15/09,A61K31/711,A61K38/00,A61K38/43,A61K45/00,A61K48/00, PC
A61P35/00,
PC A61P43/00,C07K16/32,C07K16/40,C12N1/15,C12N1/19,C12N1/21 PC
C12N5/10,C12N9/02,
PC C12P21/02,C12P21/08,G01N33/15,G01N33/50,G01N33/53/// (C12N15/09,
PC C12R1:91),
PC (C12N15/09,C12R1:19),(C12N9/02,C12R1:91),(C12P21/02,C12R1:91),
PC (C12P21/08,C12R1:91),C12N15/00,A61K37/02,A61K37/48,C12N5/00,
PC (C12N15/00,C12R1:91),(C12N15/00,C12R1:19)
CC Novel protein and DNA thereof
FH Key Location/Qualifiers
FT source 1..4955
FT Location/Qualifiers
1..4955
/organism="Homo sapiens (human)"
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

Query Match 100.0%; Score 1053; DB 6; Length 4955;
Best Local Similarity 100.0%; Pred. No. 1.1e-255;

ORIGIN

Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGCGCACCCGGAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 245 ATGGCGCACCCGGAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
Qy 61 GACACCAACGAAAGTGAATAAAGTCAAAAGAGGCCACTCCTTAAGAAAGAGTTCTCGC 120
Db 305 GACACCAACGAAAGTGAATAAAGTCAAAAGAGGCCACTCCTTAAGAAAGAGTTCTCGC 364
Qy 121 CGGTTTGTCTCTTCCAAATCCAGTACCTGATATTTGGAATAATGATAAAGCAGCAG 180
Db 365 CGGTTTGTCTCTTCCAAATCCAGTACCTGATATTTGGAATAATGATAAAGCAGCAG 424
Qy 181 GCTTCTCTTCGACAGCAGAGAGGTGACATTTATCAAGAGATCTCCCTCAGTGGAAACAG 240
Db 425 GCTTCTCTTCGACAGCAGAGAGGTGACATTTATCAAGAGATCTCCCTCAGTGGAAACAG 484
Qy 241 CTTAAGCAGATGAGAAAGTACTTCTCATCTCTCAATCTTTAGCCTTTTTCAGCCAGTGAT 300
Db 485 CTTAAGCAGATGAGAAAGTACTTCTCATCTCTCAATCTTTAGCCTTTTTCAGCCAGTGAT 544
Qy 301 GGAATTTGTAATGAAAATTTTGGTGGAGCGCTTTTAGTCAGGAGGTGCAGGTTCAGAGGCT 360
Db 545 GGAATTTGTAATGAAAATTTTGGTGGAGCGCTTTTAGTCAGGAGGTGCAGGTTCAGAGGCT 604
Qy 361 CGCTGTTTCTATGCTTTCAAATTTCTCATCAGAGATGTTCACTCAGAGATGATACAGTTTG 420
Db 605 CGCTGTTTCTATGCTTTCAAATTTCTCATCAGAGATGTTCACTCAGAGATGATACAGTTTG 664
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAATGCAATTGAA 724
Qy 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTCGCGATGATGATGAGATGAGAAA 540
Db 725 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTCGCGATGATGATGAGATGAGAAA 784
Qy 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTCTGCTGCTAGAGAGGAGTTTCTTCTCAGGA 600
Db 785 TCTACTTTTGGGGAAGAGTGGTGGCTTTCTGCTGCTAGAGAGGAGTTTCTTCTCAGGA 844
Qy 601 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 660
Db 845 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 904
Qy 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGTGCTGATGTTCCAA 720
Db 905 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGTGCTGATGTTCCAA 964
Qy 721 TACTTAGTAATAAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGTGATGCTGTCAAA 780
Db 965 TACTTAGTAATAAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGTGATGCTGTCAAA 1024
Qy 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCCAGTTCGACGTTTGAATGAATTTGCATT 840
Db 1025 ATTGAGCAGGAGTTTAAACAGAGCCTTGCCAGTTCGACGTTTGAATGAATTTGCATT 1084
Qy 841 TTGATGAACAGTACATTGAGTTTGTAGTCGACAGATTACTTGTGGAACCTTGGATTCTCA 900
Db 1085 TTGATGAACAGTACATTGAGTTTGTAGTCGACAGATTACTTGTGGAACCTTGGATTCTCA 1144
Qy 901 AAGGTTTTTCAGGAGAGAAATCCCTTTGATTTTATGGAAGAAATTTCTTTAGAGAGAAA 960
Db 1145 AAGGTTTTTCAGGAGAGAAATCCCTTTGATTTTATGGAAGAAATTTCTTTAGAGAGAAA 1204
Qy 961 ACAAATTTCTTTGAGAAACAGAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGCAGAAACC 1020
Db 1205 ACAAATTTCTTTGAGAAACAGAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGCAGAAACC 1264
Qy 1021 ACAGATAACGTCCTTTCACCTTGGATGCGAGATTTT 1053
Db 1265 ACAGATAACGTCCTTTCACCTTGGATGCGAGATTTT 1297

	RESULT 8	BD093078	Novel protein and its DNA.	4955 bp	DNA	linear	PAT 27-AUG-2002	
	LOCUS	BD093078	Novel protein and its DNA.					
	DEFINITION	BD093078	Novel protein and its DNA.					
	ACCESSION	BD093078.1	GI:22638666					
	VERSION	WO 0100799-A/2.						
	KEYWORDS	Homo sapiens (human)						
	SOURCE	Homo sapiens						
	ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.						
	REFERENCE	1 (bases 1 to 4955)						
	AUTHORS	Nakamura,Y., Arakawa,H. and Tanaka,H.						
	TITLE	Novel protein and its DNA						
	JOURNAL	Patent: WO 0100799-A 2 04-JAN-2001;						
		TAKEDA CHEMICAL INDUSTRIES LTD, YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA						
	COMMENT	OS Homo sapiens (human)						
		PN WO 0100799-A/2						
		PD 04-JAN-2001						
		PF 27-JUN-2000 WO 2000JP004189						
		PR 28-JUN-1999 JP 99P 181131, 06-JUL-1999 JP 99P 192391 PR						
		21-JAN-2000 JP 00P 017770						
		PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC						
		C12N5/04,C12N15/53,C12N1/15,C12N1/19,C12N1/21,C12N5/10 PC						
		C12P21/02,A61K38/44,						
		PC A61K48/00,C07K16/40,G01N33/53,C12Q1/26						
		CC						
	FH Key	Location/Qualifiers.						
	FEATURES	source						
		1..4955						
		/organism="Homo sapiens"						
		/mol_type="genomic DNA"						
		/db_xref="taxon:9606"						
	ORIGIN							
		Query Match	100.0%; Score 1053; DB 6; Length 4955;					
		Best Local Similarity	100.0%; Pred. No. 1.le-255;					
		Matches 1053; Conservative	0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	ATGGGCGACC CGGAAGGCGGAAAGCGGCGGTCTGGATCAGGATGAGAGATCATCTTCA	60					
DB	245	ATGGGCGACC CGGAAGGCGGAAAGCGGCGGTCTGGATCAGGATGAGAGATCATCTTCA	304					
QY	61	GACACCAACGAAGTGAAATAAAGTCAAATGAAGGCCACTCTTAAGAAAGATTCTCGC	120					
DB	305	GACACCAACGAAGTGAAATAAAGTCAAATGAAGGCCACTCTTAAGAAAGATTCTCGC	364					
QY	121	CGGTTTGTCATCTTTCCAATCCAGTAGCCCTGTATTTGGAAAATGPTATAAACAGGCACAG	180					
DB	365	CGGTTTGTCATCTTTCCAATCCAGTAGCCCTGTATTTGGAAAATGPTATAAACAGGCACAG	424					
QY	181	GCTTCCTCTCGACACGAGAAGGTGCGACTTATCAAAGGATCTCCCTCACTGGAACAAG	240					
DB	425	GCTTCCTCTCGACACGAGAAGGTGCGACTTATCAAAGGATCTCCCTCACTGGAACAAG	484					
QY	241	CTTAAGACGATGAGAAGTACTTCATCTCTCAATCTTAGCCCTTTTTTCAGCCAGTGAT	300					
DB	485	CTTAAGACGATGAGAAGTACTTCATCTCTCAATCTTAGCCCTTTTTTCAGCCAGTGAT	544					
QY	301	GGAAATGTAATGAAATTTGGTGAGCGCTTTAGTCAGGAGTGCAGGTTCAGAGGCT	360					
DB	545	GGAAATGTAATGAAATTTGGTGAGCGCTTTAGTCAGGAGTGCAGGTTCAGAGGCT	604					
QY	361	CGCTGTTCTATGGCTTTCAAAATCTCATCGAAGATGTTCACTCAGAGATGPACAGTTTG	420					
DB	605	CGCTGTTCTATGGCTTTCAAAATCTCATCGAAGATGTTCACTCAGAGATGPACAGTTTG	664					
QY	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATCAATGAA	480					
DB	665	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATCAATGAA	724					

	RESULT 8	BD093078	Novel protein and its DNA.	4955 bp	DNA	linear	PAT 27-AUG-2002	
	LOCUS	BD093078	Novel protein and its DNA.					
	DEFINITION	BD093078	Novel protein and its DNA.					
	ACCESSION	BD093078.1	GI:22638666					
	VERSION	WO 0100799-A/2.						
	KEYWORDS	Homo sapiens (human)						
	SOURCE	Homo sapiens						
	ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.						
	REFERENCE	1 (bases 1 to 4955)						
	AUTHORS	Nakamura,Y., Arakawa,H. and Tanaka,H.						
	TITLE	Novel protein and its DNA						
	JOURNAL	Patent: WO 0100799-A 2 04-JAN-2001;						
		TAKEDA CHEMICAL INDUSTRIES LTD, YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA						
	COMMENT	OS Homo sapiens (human)						
		PN WO 0100799-A/2						
		PD 04-JAN-2001						
		PF 27-JUN-2000 JP 2000JP004189						
		PR 28-JUN-1999 JP 99P 181131, 06-JUL-1999 JP 99P 192391 PR						
		21-JAN-2000 JP 00P 017770						
		PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC						
		C12N5/04,C12N15/53,C12N1/15,C12N1/19,C12N1/21,C12N5/10 PC						
		C12P21/02,A61K38/44,						
		PC A61K48/00,C07K16/40,G01N33/53,C12Q1/26						
		CC						
	FH Key	Location/Qualifiers.						
	FEATURES	source						
		1..4955						
		/organism="Homo sapiens"						
		/mol_type="genomic DNA"						
		/db_xref="taxon:9606"						
	ORIGIN							
		Query Match	100.0%; Score 1053; DB 6; Length 4955;					
		Best Local Similarity	100.0%; Pred. No. 1.le-255;					
		Matches 1053; Conservative	0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	ATGGGCGACC CGGAAGGCCGGAAGCGCGGCCTGGATCAGGATGAGAGATCATCTTCA	60					
DB	245	ATGGGCGACC CGGAAGGCCGGAAGCGCGGCCTGGATCAGGATGAGAGATCATCTTCA	304					
QY	61	GACACCAACGAAGTGAAATAAAGTCAAATGAAGACCCACTCTTAAGAAAGATTCTCGC	120					
DB	305	GACACCAACGAAGTGAAATAAAGTCAAATGAAGACCCACTCTTAAGAAAGATTCTCGC	364					
QY	121	CGGTGTGTCATCTTTCCAATCCAGTAGCCCTGATATTTGGAAAATGATATAAACAGCACAG	180					
DB	365	CGGTGTGTCATCTTTCCAATCCAGTAGCCCTGATATTTGGAAAATGATATAAACAGCACAG	424					
QY	181	GCTTCCTCTCGACACGAGAAGGTCGACTTATCAAAGGATCTCCCTCACTGGAACAAG	240					
DB	425	GCTTCCTCTCGACACGAGAAGGTCGACTTATCAAAGGATCTCCCTCACTGGAACAAG	484					
QY	241	CTTAAGACGATGAGAAGTACTTCATCTCTCACTTATAGCCTTTTTTCAGCCAGTGAT	300					
DB	485	CTTAAGACGATGAGAAGTACTTCATCTCTCACTTATAGCCTTTTTTCAGCCAGTGAT	544					
QY	301	GGAAATGTAATGAAATTTGGTGAGCGCTTTAGTCAGGAGTGCGAGGTCACAGGCT	360					
DB	545	GGAAATGTAATGAAATTTGGTGAGCGCTTTAGTCAGGAGTGCGAGGTCACAGGCT	604					
QY	361	CGCTGTTCTATGGCTTTTCAAATTTCTCATCGAAGATGTTCACTCAGAGATGPACAGTTTG	420					
DB	605	CGCTGTTCTATGGCTTTTCAAATTTCTCATCGAAGATGTTCACTCAGAGATGPACAGTTTG	664					
QY	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATCAATGAA	480					
DB	665	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATCAATGAA	724					

```
/db_xref="taxon:9606"
/chromosome="8"
/map="8q23.1"
1..4955
/gene="p53R2"
245..1300
/gene="p53R2"
/codon_start=1
/product="ribonucleotide reductase"
/protein_id="BA092434.1"
/db_xref="GI:7229086"
/translacion="MGDPPEPAAGLDQDERSSDTNESBIKNEEPLLRKSRFFVI
FPIDPIWMYKQAQASFWFAEVDLSKDPHNKIKADEKFIPIHILAFPAASDGI
TMENLVERFQVEQVPEARCFYQIILINHVSEMYSLLDITRDPKKBFLPNAIE
VMPVKKKADWALRWIADRKSTFGERVVAAVEGVFPFSGSPAALFWLKRGLMPGLT
FSNELISDGLHCDPAFLMFOVLVNPSPSERVRELIIVDAVKLEQEBLEALPVLGIG
MNCILMKQYIEFVADRLILVELGKVFQAEINPDMENISLEGKTNFFKRVSEYQRP
AVMAETDNDVFTLDADF"
4933
/polyA_site
/gene="p53R2"
/note="22 a nucleotides"
```

ORIGIN

```
Query Match 100.0%; Score 1053; DB 9; Length 4955;
Best Local Similarity 100.0%; Pred. No. 1.1e-255;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCGACCGGAAAGCGCGGAGCGCGCGGCTGGATCAGGATGAGATCATCTTCA 60
Db |
Qy 245 ATGGCGACCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGATCATCTTCA 304
Db |
Qy 61 GACACCAAGAAAGTGAATTAAGTCAATGAAGAGCCACTCTCAAGAAAGAGTTCGCG 120
Db |
Qy 305 GACACCAAGAAAGTGAATTAAGTCAATGAAGAGCCACTCTCAAGAAAGAGTTCGCG 364
Db |
Qy 121 CGGTTTGTCTTCTTCAATCCAGTACCTGATATTTGGAATATGTAACAGGCAAG 180
Db |
Qy 365 CGGTTTGTCTTCTTCAATCCAGTACCTGATATTTGGAATATGTAACAGGCAAG 424
Db |
Qy 181 GCTTCTCTTGGACAGCAGAGAGGTGCACTTATCAAGAGTCTCCCTCACTGGAACAAG 240
Db |
Qy 425 GCTTCTCTTGGACAGCAGAGAGGTGCACTTATCAAGAGTCTCCCTCACTGGAACAAG 484
Db |
Qy 241 CTTAAGCAGATGAGAAGTACTTCACTCTCACAATCTTAGCTTTTGGACGAGTAT 300
Db |
Qy 485 CTTAAGCAGATGAGAAGTACTTCACTCTCACAATCTTAGCTTTTGGACGAGTAT 544
Db |
Qy 301 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTTCCAGAGGCT 360
Db |
Qy 545 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTTCCAGAGGCT 604
Db |
Qy 361 CGCTGTTTCTATGGCTTTCAATTTCTCATCGAATGTTTCACTCAGAGATGTACAGTTTG 420
Db |
Qy 605 CGCTGTTTCTATGGCTTTCAATTTCTCATCGAATGTTTCACTCAGAGATGTACAGTTTG 664
Db |
Qy 421 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 480
Db |
Qy 665 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 724
Db |
Qy 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATGGGCGCTTGGATGGATAGCAGATAGAAAA 540
Db |
Qy 725 ACCATGCCCTATGTTAAGAAAAAAGCAGATGGGCGCTTGGATGGATAGCAGATAGAAAA 784
Db |
Qy 541 TCTACTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGGATTTTCTTCTCAGGA 600
Db |
Qy 785 TCTACTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGGATTTTCTTCTCAGGA 844
Db |
Qy 601 TCTTTTGTCTGTATTTCTGGCTTAAGAAAGAGAGTCTTATGCGCAGACTCACTTTTTC 660
Db |
Qy 845 TCTTTTGTCTGTATTTCTGGCTTAAGAAAGAGAGTCTTATGCGCAGACTCACTTTTTC 904
Db |
Qy 661 AATGAACTCATCAGCAGATGAAGGACTTCACCTGTGACTTTTGTCTGCTGATGTTCCAA 720
Db |
```

```
Db 905 AATGAACTCATCAGCAGAGATGAAGGACTTCACCTGTGACTTTTGTCTGCTGATGTTCCAA 964
Qy |
Db 721 TACTTAGTAAATAAGCCCTTCAGAAAGAGGCTCAGGAGATCATTTGTGATGCTGCTCAAA 780
Qy |
Db 965 TACTTAGTAAATAAGCCCTTCAGAAAGAGGCTCAGGAGATCATTTGTGATGCTGCTCAAA 1024
Qy |
Db 781 ATTGAGCAGGAGTATTTAAACAGAAAGCCCTTGCCAGTTGGCCTCATTTGGAATGAATTCATT 840
Db |
Db 1025 ATTGAGCAGGAGTATTTAAACAGAAAGCCCTTGCCAGTTGGCCTCATTTGGAATGAATTCATT 1084
Qy |
Qy 841 TTGATGAAAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTTGTGGAATCTTGGAATCTCA 900
Db |
Db 1085 TTGATGAAAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTTGTGGAATCTTGGAATCTCA 1144
Qy |
Qy 901 AAGGTTTTTTCAGGCGAGAAAAATCCCTTTGATTTTATGGAACAAATTTCTTTAGAAGAAAA 960
Db |
Db 1145 AAGGTTTTTTCAGGCGAGAAAAATCCCTTTGATTTTATGGAACAAATTTCTTTAGAAGAAAA 1204
Qy |
Qy 961 ACAAATTTCTTTGAGAAAAACGAGTTTTCAGAGTATCAGCGTTTTTGCAGTTATGGCAGAAACC 1020
Db |
Db 1205 ACAAATTTCTTTGAGAAAAACGAGTTTTCAGAGTATCAGCGTTTTTGCAGTTATGGCAGAAACC 1264
Qy |
Db 1021 ACAGATAACGTTCTTCACTTGGATGCGAGATTTT 1053
Db |
Db 1265 ACAGATAACGTTCTTCACTTGGATGCGAGATTTT 1297

RESULT 10
AR454877 1053 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 12 from patent US 6682917.
ACCESSION AR454877
VERSION AR454877.1 GI:42688832
KEYWORDS
SOURCE Unknown.
ORGANISM Unkown.
REFERENCE 1 (bases 1 to 1053)
AUTHORS Nakamura,Y., Arakawa,H. and Tanaka,H.
TITLE Protein having a ribonucleotide Reductase activity and a DNA
JOURNAL Patent: US 6682917-A 12 27-JAN-2004;
FEATURES Location/Qualifiers
source 1..1053
/mol_type="genomic DNA"

ORIGIN
Query Match 99.8%; Score 1051.4; DB 6; Length 1053;
Best Local Similarity 99.9%; Pred. No. 3e-255;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGGCGACCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGATCATCTTCA 60
Db |
Qy 61 GACACCAAGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTGAAGAAAGAGTTCGCG 120
Db |
Qy 61 GACACCAAGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTGAAGAAAGAGTTCGCG 120
Db |
Qy 121 CGGTTTGTCTTCTTCAATCCAGTACCTGATATTTGGAATAATGTAATAACAGGCAAG 180
Db |
Qy 121 CGGTTTGTCTTCTTCAATCCAGTACCTGATATTTGGAATAATGTAATAACAGGCAAG 180
Qy |
Qy 181 GCTTCTCTTGGACAGCAGAGAGTGCAGTTATCAAGAGTCTCCCTCACTGGAACAAG 240
Db |
Qy 181 GCTTCTCTTGGACAGCAGAGAGTGCAGTTATCAAGAGTCTCCCTCACTGGAACAAG 240
Qy |
Qy 241 CTTAAGCAGATGAGAAGTACTTCACTCTCACAATCTTAGCTTTTGGACGAGTAT 300
Db |
Qy 241 CTTAAGCAGATGAGAAGTACTTCACTCTCACAATCTTAGCTTTTGGACGAGTAT 300
Qy |
Qy 301 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTTCCAGAGGCT 360
```

Db	301	GGAAATTGTAAATGAAAAATTTGGTGGAGCGCTTTAGTCAAGAGGTGCAGGTTCCAGAGGCT	360
Qy	361	CGCTGTTTCTATGGCTTTCAAATTCATCTCAGAGAATGTTCTACTCAGAGATGTCTACAGATGTCTACAGTTTG	420
Db	361	CGCTGTTTCTATGGCTTTCAAATTCATCTCAGAGAATGTTCTACTCAGAGATGTCTACAGATGTCTACAGTTTG	420
Qy	421	CTGATAGACACTTATCATCAGAGATCCCAGAAGAAAGGGAATTTTATTTAAATGCAATTCGA	480
Db	421	CTGATAGACACTTATCATCAGAGATCCCAGAAGAAAGGGAATTTTATTTAAATGCAATTCGA	480
Qy	481	ACCATGCCCTATGTTAAGAAAAAGACAGATTGGGCCCTTGGCATGTGATAGCAGATAGAAAA	540
Db	481	ACCATGCCCTATGTTAAGAAAAAGACAGATTGGGCCCTTGGCATGTGATAGCAGATAGAAAA	540
Qy	541	TCTACTTTTGGGGAAGAGTGGTGGCCCTTTCGTCTGTCTAGAGAAGAGTTTCTTCTCAGGA	600
Db	541	TCTACTTTTGGGGAAGAGTGGTGGCCCTTTCGTCTGTCTAGAGAAGAGTTTCTTCTCAGGA	600
Qy	601	TCTTTTGTCTGTATATTCTGTGCTTAAAGAAGAGAGTCTTATGSCCAGAGACTCACATTTTTTCC	660
Db	601	TCTTTTGTCTGTATATTCTGTGCTTAAAGAAGAGAGTCTTATGSCCAGAGACTCACATTTTTTCC	660
Qy	661	AATGAACTCATCAGCAGAGATGAAGGACTTCACATGTGACTTTGCTTGCTGATGTTCCTCA	720
Db	661	AATGAACTCATCAGCAGAGATGAAGGACTTCACATGTGACTTTGCTTGCTGATGTTCCTCA	720
Qy	721	TACTTAGTAAATAAGCCTTCAGAGAAGAGGTTCAGGGAGATCATTTGTTGATGTCTCAAA	780
Db	721	TACTTAGTAAATAAGCCTTCAGAGAAGAGGTTCAGGGAGATCATTTGTTGATGTCTCAAA	780
Qy	781	ATTGAGCAGAGNTTTTAAACAGAGCCCTTGCCAGTTGGCCCTCATTTGGAATGAATTCGATT	840
Db	781	ATTGAGCAGAGNTTTTAAACAGAGCCCTTGCCAGTTGGCCCTCATTTGGAATGAATTCGATT	840
Qy	841	TTGATGAAAACAGTACATTGAGTTGTAGTCTCAGACAGATTACTTGTGGAACCTTGATTTCTCA	900
Db	841	TTGATGAAAACAGTACATTGAGTTGTAGTCTCAGACAGATTACTTGTGGAACCTTGATTTCTCA	900
Qy	901	AAGGTTTTTTCAGGCAGAAAAATCCCTTTTGAATTTATGAAAAATTTCTTTAGAGGAAAA	960
Db	901	AAGGTTTTTTCAGGCAGAAAAATCCCTTTTGAATTTATGAAAAATTTCTTTAGAGGAAAA	960
Qy	961	ACAAATTTCTTTTGAGAAAACGAGTTTTCAGAGTATCAGCGTTTTCGAGTTATGCGCAGAAACC	1020
Db	961	ACAAATTTCTTTTGAGAAAACGAGTTTTCAGAGTATCAGCGTTTTCGAGTTATGCGCAGAAACC	1020
Qy	1021	ACAGATAACGTCCTCACCTTGGATGCAGATTTT	1053
Db	1021	ACAGATAACGTCCTCACCTTGGATGCAGATTTT	1053

[illegible]

C12N15/09, A61K31/711, A61K38/00, A61K38/43, A61K45/00, A61K48/00, PC A61P35/00, A61P43/00, C07K16/32, C07K16/40, C12N1/15, C12N1/19, C12N1/21 PC C12N5/10, C12N9/02, PC C12P21/02, C12P21/08, G01N33/15, G01N33/50, G01N33/53/(C12N15/09, C12R1:91), (C12N15/09, C12R1:19), (C12N9/02, C12R1:91), (C12P21/02, C12R1:91), (C12P21/08, C12R1:91), (C12N15/00, A61K37/02, A61K37/48, C12N5/00, C12N15/00, C12R1:91), (C12N15/00, C12R1:19) CC Novel protein and DNA thereof PC Key Location/Qualifiers FH source 1..1053 FT /organism="Homo sapiens (human)".	
FEATURES	Location/Qualifiers
source	1..1053
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
ORIGIN	
Query Match	99.8%; Score 1051.4; DB 6; Length 1053;
Best Local Similarity	99.9%; Pred. No. 3e-255;
Matches 1052; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
QY 1	ATGGGGACCCGGAAAGCCGGAGCGCGCGCTGGATCAGGATGAGAGATCATCTTCA 60
Db	1 ATGGGGACCCGGAAAGCCGGAGCGCGCGCTGGATCAGGATGAGAGATCATCTTCA 60
QY 61	GACACCAACGAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGATTCTCGC 120
Db	61 GACACCAACGAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGATTCTCGC 120
QY 121	CGGTTTGTCACTTTCCAAATCCAGTACCCTGTATTTGGAAATGTATAAACAGGCACAG 180
Db	121 CGGTTTGTCACTTTCCAAATCCAGTACCCTGTATTTGGAAATGTATAAACAGGCACAG 180
QY 181	GCTTCTCTTCGACAGCAGAAGAGGTGCACTTATCAAAAGGATCTCCCTCACTGGAAACAAG 240
Db	181 GCTTCTCTTCGACAGCAGAAGAGGTGCACTTATCAAAAGGATCTCCCTCACTGGAAACAAG 240
QY 241	CTTAAGCAGATGAGAAGTACTTTCATCTCTCAATCTTAGCCCTTTTTCGAGCCAGTGAT 300
Db	241 CTTAAAGCAGATGAGAAGTACTTTCATCTCTCAATCTTAGCCCTTTTTCGAGCCAGTGAT 300
QY 301	GGAAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCT 360
Db	301 GGAATGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCT 360
QY 361	CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAAATGTTTCACTCAGAGATGTACAGTTTG 420
Db	361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAAATGTTTCACTCAGAGATGTACAGTTTG 420
QY 421	CTGATAGA CACTTACATCAGAGATCCCAAGAAAGGGAATTTTTTATTTAATGCAATTGAA 480
Db	421 CTGATAGACACTTTACATCAGAGATCCCAAGAAAGGGAATTTTTTATTTAATGCAATTGAA 480
QY 481	ACCATCCCTATCTTAAGAAAAAGCAGATTGGGCCCTTCGATGGATACGAGTAGAAAA 540
Db	481 ACCATGCCCTATGTTTAAGAAAAAGCAGATTGGGCCCTTCGATGGATACGAGTAGAAAA 540
QY 541	TCTACTTTTGGGAAAGAGTGGTGGCTTTTGCTGCTGTAGAAGGAGTTTTTCTCTCAGGA 600
Db	541 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGCTGCTGTAGAAGGAGTTTTTCTCTCAGGA 600
QY 601	TCTTTTGTGCTATATTCTGGCTTAAAGAAAGAGAGGTCTTATGCGAGACTCACTTTTTC 660
Db	601 TCTTTTGTGCTATATTCTGGCTTAAAGAAAGAGAGGTCTTATGCGAGACTCACTTTTTC 660
QY 661	AATGAACTCATCAGCAGAGATGAAGACTTCACTGTGACTTTTCTGCTTGCCTGATGTTCCAA 720
Db	661 AATGAACTCATCAGCAGAGATGAAGACTTCACTGTGACTTTTCTGCTTGCCTGATGTTCCAA 720
QY 721	TACTTAGTAAATGAAGCTTTCAGAAAGAAAGGGTCAGGAGATCATTTCTGATGCTGTCAA 780

```
Db 721 TACTTAGTAATAAGCCTTCAGAGAAAGGCTCAGGAGATCAATTGTTGCTGTCAAA 780
Qy 781 ATTGAGCAGGAGTTTAAACAGAGAGCCTTGCCAGTTGSCCTCAATTGGAATGAATTGCATT 840
Db 781 ATTGAGCAGGAGTTTAAACAGAGAGCCTTGCCAGTTGSCCTCAATTGGAATGAATTGCATT 840
Qy 841 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
Db 841 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
Qy 901 AAGGTTTTTCAGGAGAGAAATCCTTTTGATTTTATGGAATGAATTCTTTAGAGAGAAA 960
Db 901 AAGGTTTTTCAGGAGAGAAATCCTTTTGATTTTATGGAATGAATTCTTTAGAGAGAAA 960
Qy 961 ACAAATTTCTTTGAGAAACGAGTTTCAGAGATCAGCGTTTTCAGATTATGCGCAGAAACC 1020
Db 961 ACAAATTTCTTTGAGAAACGAGTTTCAGAGATCAGCGTTTTCAGATTATGCGCAGAAACC 1020
Qy 1021 ACAGATAACGCTCTTCAACCTTGGATGCAGATTTT 1053
Db 1021 ACAGATAACGCTCTTCAACCTTGGATGCAGATTTT 1053

RESULT 12
LOCUS BD093086 1053 bp DNA linear PAT 27-AUG-2002
DEFINITION Novel protein and its DNA.
ACCESSION BD093086
VERSION BD093086.1 GI:22638674
KEYWORDS WO 0100799-A/10.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1053)
AUTHORS Nakamura,Y., Arakawa,H. and Tanaka,H.
TITLE Novel protein and its DNA
JOURNAL Patent: WO 0100799-A 10 04-JAN-2001;
TAKEDA CHEMICAL INDUSTRIES LTD, YUSUKE NAKAMURA, HIROFUMI ARAKAWA,
HIROSHI TANAKA
COMMENT OS Homo sapiens (human)
PN WO 0100799-A/10
PD 04-JAN-2001
PF 27-JUN-2000 WO 2000JP004189
PR 28-JUN-1999 JP 99P 181131,06-JUL-1999 JP 99P 192391 PR
21-JAN-2000 JP 00P 017770
PI YUSUKE NAKAMURA,HIROFUMI ARAKAWA,HIROSHI TANAKA PC
C12N9/04,C12N15/53,C12N1/15,C12N1/19,C12N1/21,C12N5/10 PC
,C12P21/02,A61K38/44,
PC A61K48/00,C07K16/40,G01N33/53,C12Q1/26
CC

FEATURES
source location/Qualifiers
1..1053
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 99.8%; Score 1051.4; DB 6; Length 1053;
Best Local Similarity 99.9%; Pred. No. 3e-255;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGGCGCACCGGAAAGCCGGAAGCGCGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 ATGGCGCACCGGAAAGCCGGAAGCGCGCTGGATCAGGATGAGAGATCATCTTCA 60
Qy 61 GACACCAACGAACTGAAATGAATCAATGAAGCCACTCCCTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAACGAACTGAAATGAATCAATGAAGCCACTCCCTAAGAAAGAGTTCTCGC 120
Qy 121 CGGTTTGTCTCTTTCCAAATCCAGTACCCTCGATATTTGGAAATGTATAAACAGGCACAG 180
```

```
Db 121 CGGTTTGTCTCTTTCCAAATCCAGTACCCTCGATATTTGGAAATGTATAAACAGGCACAG 180
Qy 181 GCTTCCCTCTCGACAGCAGAGAGAGTGCACCTTATCABAAGGATCTCCCTCACTGGACACAG 240
Db 181 GCTTCCCTCTCGACAGCAGAGAGAGTGCACCTTATCABAAGGATCTCCCTCACTGGACACAG 240
Qy 241 CTTAAAGCAGATGAGAGTACTTTCATCTCTCACTCTTAGCCTTTTTCAGCCAGTGTAT 300
Db 241 CTTAAAGCAGATGAGAGTACTTTCATCTCTCACTCTTAGCCTTTTTCAGCCAGTGTAT 300
Qy 301 GGAATGTGAAATGAAATTTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
Db 301 GGAATGTGAAATGAAATTTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
Qy 361 CGCTGTTTCTATGCTCTTCAAAATTTCTCATCGAGATGTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGCTCTTCAAAATTTCTCATCGAGATGTTCACTCAGAGATGTACAGTTTG 420
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAAATCAATTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAAATCAATTGAA 480
Qy 481 ACCATGCCCTATGTTAAAGAAAACAGATTGGCGCTTTCGATGGATGATAGATAGAAA 540
Db 481 ACCATGCCCTATGTTAAAGAAAACAGATTGGCGCTTTCGATGGATGATAGATAGAAA 540
Qy 541 TCTACTTTTGGGAAAGAGTGTGGCTTTCGCTGCTGTAGAAGGAGTTTCTTCTCAGA 600
Db 541 TCTACTTTTGGGAAAGAGTGTGGCTTTCGCTGCTGTAGAAGGAGTTTCTTCTCAGA 600
Qy 601 TCTTTTGTGCTATATTCTGGCTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
Db 601 TCTTTTGTGCTATATTCTGGCTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
Qy 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTGTATGTTCCA 720
Db 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTGTATGTTCCA 720
Qy 721 TACTTAGTAATAAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTATGCTGTCAAA 780
Db 721 TACTTAGTAATAAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTATGCTGTCAAA 780
Qy 781 ATTGAGCAGGAGTTTAAACAGAGAGCCTTGCCAGTTTGGCCTCATTTGCTGTATGCTCA 840
Db 781 ATTGAGCAGGAGTTTAAACAGAGAGCCTTGCCAGTTTGGCCTCATTTGCTGTATGCTCA 840
Qy 841 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
Db 841 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
Qy 901 AAGGTTTTTCAGGAGAGAAATCCTTTTGATTTTATGGAATGAATTCTTTAGAGAGAAA 960
Db 901 AAGGTTTTTCAGGAGAGAAATCCTTTTGATTTTATGGAATGAATTCTTTAGAGAGAAA 960
Qy 961 ACAAATTTCTTTGAGAAACGAGTTTCAGAGATCAGCGTTTTCAGATTATGCGCAGAAACC 1020
Db 961 ACAAATTTCTTTGAGAAACGAGTTTCAGAGATCAGCGTTTTCAGATTATGCGCAGAAACC 1020
Qy 1021 ACAGATAACGCTCTTCAACCTTGGATGCAGATTTT 1053
Db 1021 ACAGATAACGCTCTTCAACCTTGGATGCAGATTTT 1053

RESULT 13
LOCUS AR454870 1081 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 4 from patent US 6682917.
ACCESSION AR454870
VERSION AR454870.1 GI:42688825
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
```


181 GCTTCTCTCTGGACAGCAGAGAGGTGCGACTTATCAAGAGATCTCCCTCACTGGAACAAG 240
Db |||||
200 GCTTCTCTCTGGACAGCAGAGAGGTGCGACTTATCAAGAGATCTCCCTCACTGGAACAAG 259
Qy |||||
241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACATCTTAGCGCTTTTGGACGAGTGAT 300
Db |||||
260 CTTAAAGCAGATGAGAAGTACTTCACTCTCACATCTTAGCGCTTTTGGACGAGTGAT 319
Qy |||||
301 GGAATCTGTAATGAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCGAGTCTCCAGAGGCT 360
Db |||||
320 GGAATCTGTAATGAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCGAGTCTCCAGAGGCT 379
Qy |||||
361 CGCTGTTTCTATGCTTTCAAAATCTCATCAGAGATGTTCACTCAGAGATGTCAGATTTG 420
Db |||||
380 CGCTGTTTCTATGCTTTCAAATCTCATCAGAGATGTTCACTCAGAGATGTCAGATTTG 439
Qy |||||
421 CTGATAGACATTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTGAA 480
Db |||||
440 CTGATAGACATTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTGAA 499
Qy |||||
481 ACCATGCCCTATGTTAAGAAAAGAGCAGATTTGGCGCTTGGATGATAGCAGATAGAAA 540
Db |||||
500 ACCATGCCCTATGTTAAGAAAAGAGCAGATTTGGCGCTTGGATGATAGCAGATAGAAA 559
Qy |||||
541 TCTACTTTTGGGAAAAGAGTGGTGGCTTTGCTGCTGTAGAGGAGTCTTCTCTCAGGA 600
Db |||||
560 TCTACTTTTGGGAAAAGAGTGGTGGCTTTGCTGCTGTAGAGGAGTCTTCTCTCAGGA 619
Qy |||||
601 TCTTTTGGCTGTAATTTCTGGCTAAAGAGAGAGGTCTTATGCGAGACCTCACTTTTTC 660
Db |||||
620 TCTTTTGGCTGTAATTTCTGGCTAAAGAGAGAGGTCTTATGCGAGACCTCACTTTTTC 679
Qy |||||
661 AATGAACCTCATCAGCAGATGAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db |||||
680 AATGAACCTCATCAGCAGATGAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 739
Qy |||||
721 TACTTAGTAAATAGCCCTTCAGAGAAAAGGTGAGGAGATCATTTGTTGATGCTGTCAA 780
Db |||||
740 TACTTAGTAAATAGCCCTTCAGAGAAAAGGTGAGGAGATCATTTGTTGATGCTGTCAA 799
Qy |||||
781 ATTGAGCAGGAGTTTAAACAGAGCCCTGCGAGTTGCGCTCAITGGAATGAAATTCATT 840
Db |||||
800 ATTGAGCAGGAGTTTAAACAGAGCCCTGCGAGTTGCGCTCAITGGAATGAAATTCATT 859
Qy |||||
841 TTGATGAAACAGTACATTTGATGCTGACAGATTAATCTGTGGAACCTTGGATTTCTCA 900
Db |||||
860 TTGATGAAACAGTACATTTGATGCTGACAGATTAATCTGTGGAACCTTGGATTTCTCA 919
Qy |||||
901 AAGGTTTTTTCAGGCAGAAAATCCCTTTGATTTTATGAAAACATTTCTTTAGAGGAAAA 960
Db |||||
920 AAGGTTTTTTCAGGCAGAAAATCCCTTTGATTTTATGAAAACATTTCTTTAGAGGAAAA 979
Qy |||||
961 ACRAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC 1020
Db |||||
980 ACRAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC 1039
Qy |||||
1021 ACAGATTAAGCTTTCACTCTGGATGCGAGATTTT 1053
Db |||||
1040 ACAGATAACGTCTTTCACTCTGGATGCGAGATTTT 1072

RESULT 15
BD093079
LOCUS BD093079 1081 bp DNA linear PAT 27-AUG-2002
DEFINITION Novel protein and its DNA.
ACCESSION BD093079
VERSION BD093079.1 GI:22638667
KEYWORDS WO 0100799-A/3.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1081)

AUTHORS Nakamura,Y., Arakawa,H. and Tanaka,H.
TITLE Novel protein and its DNA
JOURNAL Patent: WO 0100799-A 3 04-JAN-2001;
TAKEDA CHEMICAL INDUSTRIES LTD,YOSUKE NAKAMURA,HIROFUMI ARAKAWA,
HIROSHI TANAKA
COMMENT OS Homo sapiens (human)
PN WO 0100799-A/3
PD 04-JAN-2001
PF 27-JUN-2000 WO 2000JP004189
PR 28-JUN-1999 JP 99P 181131,06-JUL-1999 JP 99P 192391 PR
21-JAN-2000 JP 00P 017770
PI YOSUKE NAKAMURA,HIROFUMI ARAKAWA,HIROSHI TANAKA PC
C12N9/04,C12N15/53,C12N1/15,C12N1/19,C12N1/21,C12N5/10 PC
C12P21/02,A61K38/44,
PC A61K48/00,C07K16/40,G01N33/53,C12Q1/26
CC
FH Key Location/Qualifiers.
FEATURES
source 1..1081
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.8%; Score 1051.4; DB 6; Length 1081;
Best Local Similarity 99.9%; Pred. No. 3e-255;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGGCGCACCGGAAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db |||||
20 ATGGCGCACCGGAAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 79
Qy 61 GACACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGGTTCTCGC 120
Db |||||
80 GACACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGGTTCTCGC 139
Qy 121 CGGTTTGTCTATCTTTTCAATCCAGTACCTCATATTTGGAAAATCTATAAACAGCACAG 180
Db |||||
140 CGGTTTGTCTATCTTTCCATCCAGTACCTCATATTTGGAAAATCTATAAACAGCACAG 199
Qy 181 GCTTCTCTTCTGGACAGCAGAGAGGTGCGATTTATCAAAAGGATCTCCCTCACTGGAAACAAG 240
Db |||||
200 GCTTCTCTTCTGGACAGCAGAGAGGTGCGATTTATCAAAAGGATCTCCCTCACTGGAAACAAG 259
Qy 241 CTTAAACAGATGAGAAGTACTTCACTCTCACATCTTAGCGCTTTTTCAGGCGAGTGAT 300
Db |||||
260 CTTAAACAGATGAGAAGTACTTCACTCTCACATCTTAGCGCTTTTTCAGGCGAGTGAT 319
Qy 301 GGAATCTGTAATGAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCGAGTCTCCAGAGGCT 360
Db |||||
320 GGAATCTGTAATGAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCGAGTCTCCAGAGGCT 379
Qy 361 CGCTGTTTCTATGCGCTTTCAAATTTCTCATCAGAGATGTTTCACTCAGAGATGTCAGATTTG 420
Db |||||
380 CGCTGTTTCTATGCGCTTTCAAATTTCTCATCAGAGATGTTTCACTCAGAGATGTCAGATTTG 439
Qy 421 CTGATAGACATTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTGAA 480
Db |||||
440 CTGATAGACATTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTGAA 499
Qy 481 ACCATGCCCTATGTTAAGAAAAGAGCAGATTTGGCGCTTGGATGATAGCAGATAGAAA 540
Db |||||
500 ACCATGCCCTATGTTAAGAAAAGAGCAGATTTGGCGCTTGGATGATAGCAGATAGAAA 559
Qy 541 TCTACTTTTGGGAAAAGAGTGGTGGCTTTGCTGCTGTAGAGGAGTCTTCTCTCAGGA 600
Db |||||
560 TCTACTTTTGGGAAAAGAGTGGTGGCTTTGCTGCTGTAGAGGAGTCTTCTCTCAGGA 619
Qy 601 TCTTTTGGCTGTAATTTCTGGCTAAAGAGAGAGGTCTTATGCGAGACCTCACTTTTTC 660
Db |||||
620 TCTTTTGGCTGTAATTTCTGGCTAAAGAGAGAGGTCTTATGCGAGACCTCACTTTTTC 679
Qy 661 AATGAACCTCATCAGCAGATGAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720

Db		680	AATGAACCTCATCAGCAGAGATGAAGGACTTCACCTGTGACTTTGCTTGCTGATGTCCAA	739
Qy		721	TACTTAGTAATAAGCCTTTCAGAAAGAAAGGCTCAGGGAGATCATTTGTTGATGCTGTCAAA	780
Db		740	TACTTAGTAATAAGCCTTTCAGAAAGAAAGGCTCAGGGAGATCATTTGTTGATGCTGTCAAA	799
Qy		781	ATTGAGCAGGAGTTTAAACAGAAAGCCTTGCCAGTTGGCCTCATTTGGAATGAATTGCATT	840
Db		800	ATTGAGCAGGAGTTTAAACAGAAAGCCTTGCCAGTTGGCCTCATTTGGAATGAATTGCATT	859
Qy		841	TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTACTTTGGAACTTTGGATTCTCA	900
Db		860	TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTACTTTGGAACTTTGGATTCTCA	919
Qy		901	AAGGTTTTTCAGGCAGAAAAATCCCTTTTGAATTTTATGGAACATTTCTTTAGNAGGAAAA	960
Db		920	AAGGTTTTTCAGGCAGAAAAATCCCTTTTGAATTTTATGGAACATTTCTTTAGNAGGAAAA	979
Qy		961	ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTTGCAAGTTATGGCAGAAACC	1020
Db		980	ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTTGCAAGTTATGGCAGAAACC	1039
Qy		1021	ACAGATAACGTTCTTCACCTTGGATGCAGATTTT	1053
Db		1040	ACAGATAACGTTCTTCACCTTGGATGCAGATTTT	1072

Search completed: September 26, 2005, 01:35:17
 Job time : 3285 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 25, 2005, 18:03:30 ; Search time 463 Seconds
(without alignments)
13463.256 Million cell updates/sec

Title: US-10-698-228-2

Perfect score: 1053

Sequence: 1 atggcgaccgcgaaagcc.....tcaccttgatgcagatttt 1053

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N Geneseq_16Dec04.*

1: Geneseqn1980s.*

2: Geneseqn1990s.*

3: Geneseqn2000s.*

4: Geneseqn2001as.*

5: Geneseqn2001bs.*

6: Geneseqn2002as.*

7: Geneseqn2002bs.*

8: Geneseqn2003as.*

9: Geneseqn2003bs.*

10: Geneseqn2003cs.*

11: Geneseqn2003ds.*

12: Geneseqn2004as.*

13: Geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1053	100.0	1053	4	Aaf32438 Human rib
2	1053	100.0	2596	3	Aa12411 CDNA enco
3	1053	100.0	4955	4	Aaf32439 Human rib
4	1053	100.0	4955	13	ADR24210 Breast ca
5	1051.4	99.8	1053	4	Aaf32447 Human rib
6	1051.4	99.8	1081	4	Aaf32440 Human rib
7	1049.8	99.7	1601	4	Aa14924 Human cDN
8	650.6	61.8	706	13	Adq57092 Novel can
9	578.4	54.9	1989	3	Aac78111 Human can
10	578.4	54.9	2215	12	Adk70302 Respirato
11	578.4	54.9	2216	10	Adj56536 Murine cD
12	578.4	54.9	2482	4	Aa544917 Human con
13	578.4	54.9	2500	6	ABL65414 Lung can
14	578.4	54.9	2500	6	ABL65517 Lung can
15	578.4	54.9	2500	6	ABL65859 Lung can
16	578.4	54.9	2500	8	ABX10335 DNA enco
17	578.4	54.9	2500	11	AD132044 Human cDN
18	578.4	54.9	2500	12	ADN04443 Antipsoi
19	578.4	54.9	2500	12	Adq09273 Human RRM
20	578.4	54.9	2500	13	ACN37637 Tumour-as

21	578.4	54.9	3294	4	AAH73225 Human cer
22	578.4	54.9	3393	12	ADN03788 Antipsoi
23	578.4	54.9	3393	12	ADN03788 Human PRO
24	569	54.0	2641	4	AAS44745 Human ful
25	568.8	54.0	1170	2	AAV05641 Human rib
26	561	53.3	1328	12	ADO57308 DNA enco
27	547.8	52.0	978	4	AAS41006 cDNA enco
28	545.4	51.8	977	6	ABL90228 Human pol
29	528.8	50.2	608	4	AAH07707 Human cDN
30	514	48.8	1371	5	AAH79474 DNA enco
31	463	44.0	1289	4	ABL14627 Drosophil
32	442	42.0	1146	6	AAS62387 cDNA sequ
33	441	41.9	481	2	AAH51874 Human sec
34	432.2	41.0	1218	8	AAH20895 Aspergill
35	429.8	40.8	14176	2	AAT84564 Swinepox
36	429.8	40.8	14176	4	Aaf84949 Reverse c
37	429.8	40.8	14176	4	Aaf84948 Nucleotid
38	420.4	39.9	3945	4	ABL14626 Drosophil
39	403.2	38.3	963	13	ADS58378 Bacteri
40	397	37.7	1146	8	ABT19075 Aspergill
41	392.6	37.3	1173	13	ADT47814 Bacteri
42	392	37.2	1242	6	ABZ32250 Candida a
43	386.6	36.7	1292	13	ADS49379 Bacteri
44	380.4	36.1	1200	10	ACC61238 Gene sequ
45	380.4	36.1	1200	10	ADK63241 Disease t

ALIGNMENTS

RESULT 1

AAF32438
ID AAF32438 standard; cDNA; 1053 BP.

XX AAF32438;

XX 18-APR-2001 (first entry)

XX Human ribonucleotide reductase TP53R2H nucleotide sequence SEQ ID NO:2.

XX Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.

XX Homo sapiens.

XX WO200100799-A1.

XX 04-JAN-2001.

XX 27-JUN-2000; 2000WO-JP004189.

XX 28-JUN-1999; 99JP-00181131.

XX 06-JUL-1999; 99JP-00192391.

XX 21-JAN-2000; 2000JP-00017770.

XX (TAKE) TAKEDA CHEM IND LTD.

XX (NAKA/) NAKAMURA Y.

XX Nakamura Y, Arakawa H, Tanaka H;

XX WPI; 2001-112446/12.

XX P-PSDB; AAB69050.

XX Ribonucleotide reductase involved in DNA repair and DNA encoding it, for diagnosis, treatment and prevention of cancer.

XX Claim 5; Fig 1-3; 102pp; Japanese.

XX The present sequence encodes a human ribonucleotide reductase designated TP53R2H. TP53R2H has cytostatic activity. TP53R2H is part of the DNA repair mechanism and its activity is induced by p53. It can be used for the treatment, prevention and diagnosis of a wide range of cancers

XX Sequence 1053 BP; 314 A; 192 C; 246 G; 301 T; 0 U; 0 Other;

Query Match	100.0%;	Score 1053;	DB 4;	Length 1053;
Best Local Similarity	100.0%;	Pred. No. 1e-290;		
Matches 1053;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGGGCGACCGCGAAAGCGCGGCGCGGCTGGATCAGATGAGAGATCATCTTCA	60	
DB	1	ATGGGCGACCGCGAAAGCGCGGCGGCTGGATCAGATGAGAGATCATCTTCA	60	
QY	61	GACACCAACGAAAGTGAATAAATGTAAGAGGCCACTCCTTAAGAAAGAGTTCTCGC	120	
DB	61	GACACCAACGAAAGTGAATAAATGTAAGAGGCCACTCCTTAAGAAAGAGTTCTCGC	120	
QY	121	CGGTTTGTCATCTTTTCCAATCCAGTACCTCGATATTTTGAAATATGTAACACGACACAG	180	
DB	121	CGGTTTGTCATCTTTTCCAATCCAGTACCTCGATATTTTGAAATATGTAACACGACACAG	180	
QY	181	GCCTTCCTTCGACAGCAGAAAGAGTCGACTTATCAAAAGGATCTCCCTCACTCGAAACAAG	240	
DB	181	GCCTTCCTTCGACAGCAGAAAGAGTCGACTTATCAAAAGGATCTCCCTCACTCGAAACAAG	240	
QY	241	CTTAAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTAGCCCTTTTTCAGCCAGCTGAT	300	
DB	241	CTTAAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTAGCCCTTTTTCAGCCAGCTGAT	300	
QY	301	GGAAATGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGTTCCAGAGGCT	360	
DB	301	GGAAATGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGTTCCAGAGGCT	360	
QY	361	CGCTGTTTCTATCGCTTTCAAAATCTCATCGAATGTTTCACTCAGAGATGTACAGTTTG	420	
DB	361	CGCTGTTTCTATCGCTTTCAAAATCTCATCGAATGTTTCACTCAGAGATGTACAGTTTG	420	
QY	421	CTGATAGACACTTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATTTGAA	480	
DB	421	CTGATAGACACTTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATTTGAA	480	
QY	481	ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCCCTTGCATGGAATAGAGATAAGAAA	540	
DB	481	ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCCCTTGCATGGAATAGAGATAAGAAA	540	
QY	541	TCTACTTTTGGGAAAGAGTGGGCTTTGCTGCTGTAGAGAGTTTCTTCTCAGGA	600	
DB	541	TCTACTTTTGGGAAAGAGTGGGCTTTGCTGCTGTAGAGAGTTTCTTCTCAGGA	600	
QY	601	TCCTTTTGCTGCTATATTTCTGGCTTAAAGACAGAGTCTTATGCCAGACATCCTTTTTTCC	660	
DB	601	TCCTTTTGCTGCTATATTTCTGGCTTAAAGACAGAGTCTTATGCCAGACATCCTTTTTTCC	660	
QY	661	AATGAACCTCATCAGCAGATGAAGGACTTCACTGTGACTTTGCTGGCTGATGTTCCAA	720	
DB	661	AATGAACCTCATCAGCAGATGAAGGACTTCACTGTGACTTTGCTGGCTGATGTTCCAA	720	
QY	721	TACTTAGTAATTAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCAAA	780	
DB	721	TACTTAGTAATTAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCAAA	780	
QY	781	ATTGAGCAGAGTTTTTAAACAGAGCCTTCAGGTTGGCCTCATTTGGAATGAATTCATTT	840	
DB	781	ATTGAGCAGAGTTTTTAAACAGAGCCTTCAGGTTGGCCTCATTTGGAATGAATTCATTT	840	
QY	841	TTGATGAAAACAGTACATTTGATGTTAGCTGACAGATTTACTTGTGGAATCTGGATTTCTCA	900	
DB	841	TTGATGAAAACAGTACATTTGATGTTAGCTGACAGATTTACTTGTGGAATCTGGATTTCTCA	900	
QY	901	AAGGTTTTTTCAGCGAGAAAATCCTTTTGTATTTTATGGAACAACTTCTTTAGAGGAAAA	960	
DB	901	AAGGTTTTTTCAGCGAGAAAATCCTTTTGTATTTTATGGAACAACTTCTTTAGAGGAAAA	960	
QY	961	ACAAATTTCTTTGAGAAAACGAGTTTCAGAGTATCAGGCTTTTCAGTATGCGCAAAACC	1020	
DB	961	ACAAATTTCTTTGAGAAAACGAGTTTCAGAGTATCAGGCTTTTCAGTATGCGCAAAACC	1020	

```
Db 56 ATGGCGGACCCGGAAGCCGGAAGCCGGCTGGATCAGGATGAGAGATCATCTTCA 115
Qy 61 GACACCAACGAAAGTGAATAAGTCAAAATGAAGAGCCACTCCTAAGAAAGAGTTCTCGC 120
Db 116 GACACCAACGAAAGTGAATAAGTCAAAATGAAGAGCCACTCCTAAGAAAGAGTTCTCGC 175
Qy 121 CGTTTGTGTCATCTTCCAAATCCAGTACCTGATATTTGGAAATGTATPAAACAGCACAG 180
Db 176 CGTTTGTGTCATCTTCCAAATCCAGTACCTGATATTTGGAAATGTATPAAACAGCACAG 235
Qy 181 GCTTCCTTCTGGACAGCAGAGAGGTGCACTTATCAAAAGGATCTCCCTCAGTGGAAACAAG 240
Db 236 GCTTCCTTCTGGACAGCAGAGAGGTGCACTTATCAAAAGGATCTCCCTCAGTGGAAACAAG 295
Qy 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACATCTTAGCCTTTTGGCAGCCAGTGAT 300
Db 296 CTTAAAGCAGATGAGAAGTACTTCACTCTCACATCTTAGCCTTTTGGCAGCCAGTGAT 355
Qy 301 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAAGAGGTGCGAGTTCCAGAGGCT 360
Db 356 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAAGAGGTGCGAGTTCCAGAGGCT 415
Qy 361 CGCTGTTCTATGCGTTTCAAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db 416 CGCTGTTCTATGCGTTTCAAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 475
Qy 421 CTGATAGACACTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 480
Db 476 CTGATAGACACTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 535
Qy 481 ACCATGCCCTATGTTAAGAAAAGAGAGATGGGCTTGGATGATAGCAGATAGAAA 540
Db 536 ACCATGCCCTATGTTAAGAAAAGAGAGATGGGCTTGGATGATAGCAGATAGAAA 595
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCTTCTGCTGTAGAGAGTGTCTTCTCAGGA 600
Db 596 TCTACTTTTGGGAAAGAGTGGTGGCTTCTGCTGTAGAGAGTGTCTTCTCAGGA 655
Qy 601 TCTTTTGTGCTATATTTCTGCTTAAAGAGAGAGTCTTATGCCAGGACTCATTCTTCC 660
Db 656 TCTTTTGTGCTATATTTCTGCTTAAAGAGAGAGTCTTATGCCAGGACTCATTCTTCC 715
Qy 661 AATGAACCTCATCAGCAGATGAAAGGCTTCACTGTGACTTTGCTTGGCTGATGTCAA 720
Db 716 AATGAACCTCATCAGCAGATGAAAGGCTTCACTGTGACTTTGCTTGGCTGATGTCAA 775
Qy 721 TACTTAGTAAATAAGCCTTTCAGAAAGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAA 780
Db 776 TACTTAGTAAATAAGCCTTTCAGAAAGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAA 835
Qy 781 ATTGAGCGAGGATTTTAAACAGAACCTTGGCCAGTGGCCCTCATTTGGAATGAATTCATT 840
Db 836 ATTGAGCGAGGATTTTAAACAGAACCTTGGCCAGTGGCCCTCATTTGGAATGAATTCATT 895
Qy 841 TTGATGAAACAGTACATTTGAGTTGTAGTACACAGATTTACTTTGGAACCTTGGATTTCTCA 900
Db 896 TTGATGAAACAGTACATTTGAGTTGTAGTACACAGATTTACTTTGGAACCTTGGATTTCTCA 955
Qy 901 AAGGTTTTTTCAGGCAGAAAATCCTTTTGAATTTATGGAACAATTTCTTTAGAAAGGAAA 960
Db 956 AAGGTTTTTTCAGGCAGAAAATCCTTTTGAATTTATGGAACAATTTCTTTAGAAAGGAAA 1015
Qy 961 ACAAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020
Db 1016 ACAAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1075
Qy 1021 ACAGATAACGCTTTCACCTTGGATGCGAGTTTTT 1053
Db 1076 ACAGATAACGCTTTCACCTTGGATGCGAGTTTTT 1108
```

RESULT 3

AAF32439

ID AAF32439 standard; cDNA; 4955 BP.

XX AC AAF32439;

XX DT 18-APR-2001 (first entry)

XX DE Human ribonucleotide reductase related nucleotide sequence SEQ ID NO:3.

XX KW Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.

XX OS Homo sapiens.

XX PN WO200100799-A1.

XX PD 04-JAN-2001.

XX PF 27-JUN-2000; 2000WO-JP004189.

XX PR 28-JUN-1999; 99JP-00181131.

XX PR 06-JUL-1999; 99JP-00192391.

XX PR 21-JAN-2000; 2000JP-00017770.

XX PA (TAKB) TAKEDA CHEM IND LTD.

XX PA (NAKA/) NAKAMURA Y.

XX PI Nakamura Y, Arakawa H, Tanaka H;

XX XX WPI; 2001-112446/12.

XX PT Ribonucleotide reductase involved in DNA repair and DNA encoding it, for

XX PS diagnosis, treatment and prevention of cancer.

XX PS Example 2; Page 87-90; 102pp; Japanese.

XX CC The present invention describes a human ribonucleotide reductase

XX CC designated TP53R2H. TP53R2H has cytostatic activity. TP53R2H is part of

XX CC the DNA repair mechanism and its activity is induced by p53. It can be

XX CC used for the treatment, prevention and diagnosis of a wide range of

XX CC cancers. The present sequence represents a human ribonucleotide reductase

XX CC related sequence which is used in an example from the present invention

XX SQ Sequence 4955 BP; 1526 A; 803 C; 999 G; 1627 T; 0 U; 0 Other;

Query Match 100.0%; Score 1053; DB 4; Length 4955;

Best Local Similarity 100.0%; Pred. No. 2.1e-290;

Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCGGACCCGGAAGCCGGAAGCCGGCTGGATCAGGATGAGAGATCATCTTCA 60

Db 245 ATGGCGGACCCGGAAGCCGGAAGCCGGCTGGATCAGGATGAGAGATCATCTTCA 304

Qy 61 GACACCAACGAAAGTGAATAAGTCAAAATGAAGAGCCACTCCTAAGAAAGAGTTCTCGC 120

Db 305 GACACCAACGAAAGTGAATAAGTCAAAATGAAGAGCCACTCCTAAGAAAGAGTTCTCGC 364

Qy 121 CGTTTGTGTCATCTTCCAAATCCAGTACCTGATATTTGGAAATGTATPAAACAGCACAG 180

Db 365 CGTTTGTGTCATCTTCCAAATCCAGTACCTGATATTTGGAAATGTATPAAACAGCACAG 424

Qy 181 GCTTCCTTCTGGACAGCAGAGAGGTGCACTTATCAAAAGGATCTCCCTCAGTGGAAACAAG 240

Db 425 GCTTCCTTCTGGACAGCAGAGAGGTGCACTTATCAAAAGGATCTCCCTCAGTGGAAACAAG 484

Qy 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACATCTTAGCCCTTTTTCAGCCAGTGAT 300

Db 485 CTTAAAGCAGATGAGAAGTACTTCACTCTCACATCTTAGCCCTTTTTCAGCCAGTGAT 544

Qy 301 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAAGAGGTGCGAGTTCCAGAGGCT 360

Db 545 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAAGAGGTGCGAGTTCCAGAGGCT 604

Qy 361 CGCTGTTCTATGCGTTTCAAAATTTCTCATCGAGATGTTTCACTCAGAGATGTACAGTTTG 420

Db 605 CGCTGTTCTATGCGCTTTCAAAATCTCATCAGAAATGTTCACTCAGAGATGTACAGTTTG 664
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 480
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 724
Qy 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCCCTTGCGATGATAGCAGATAGAAAA 540
Db 725 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCCCTTGCGATGATAGCAGATAGAAAA 784
Qy 541 TCTACTTTTGGGGAAGAGTGGTGGCCCTTGCTGCTGTGAAGAGTGTCTTCTTCAGGA 600
Db 785 TCTACTTTTGGGGAAGAGTGGTGGCCCTTGCTGCTGTGAAGAGTGTCTTCTTCAGGA 844
Qy 601 TCTTTTGTCTGTATTTCTGCTTAAGAAAGAGAGTGTCTTATGCCAGACACCTTTTTC 660
Db 845 TCTTTTGTCTGTATTTCTGCTTAAGAAAGAGAGTGTCTTATGCCAGACACCTTTTTC 904
Qy 661 AATGAACCTCATCAGCAGATGAAGGACTTCACCTGTGACTTTTGTCTGCTGATGTTCCAA 720
Db 905 AATGAACCTCATCAGCAGATGAAGGACTTCACCTGTGACTTTTGTCTGCTGATGTTCCAA 964
Qy 721 TACTTAGTAATAAGCCTTCAGAAAGAGGTGTCAGGAGATCAATGTTGATGCTGTCAAA 780
Db 965 TACTTAGTAATAAGCCTTCAGAAAGAGGTGTCAGGAGATCAATGTTGATGCTGTCAAA 1024
Qy 781 ATTGAGCAGAGTGTAAACAGAGAGCCTTGCAGTGGCCTCATTTGGAATGATTCATT 840
Db 1025 ATTGAGCAGAGTGTAAACAGAGAGCCTTGCAGTGGCCTCATTTGGAATGATTCATT 1084
Qy 841 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTAATCTTGTGAACTTTGATTTCTCA 900
Db 1085 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTAATCTTGTGAACTTTGATTTCTCA 1144
Qy 901 AAGTGTTCAGGAGAAAAATCCCTTTTGAATTTATGAAAAACAATTTCTTTAGAGGAAAA 960
Db 1145 AAGTGTTCAGGAGAAAAATCCCTTTTGAATTTATGAAAAACAATTTCTTTAGAGGAAAA 1204
Qy 961 ACAAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGGGTTTTCAGTATGCGAGAAC 1020
Db 1205 ACAAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGGGTTTTCAGTATGCGAGAAC 1264
Qy 1021 ACAGATAACCTCTTTCACCTTTGGATGACAGATTTT 1053
Db 1265 ACAGATAACCTCTTTCACCTTTGGATGACAGATTTT 1297

RESULT 4
ID ADR24210
XX ADR24210 standard; DNA; 4955 BP.
AC ADR24210;
XX
XX
DT 21-OCT-2004 (first entry)
XX
XX Breast cancer prognosis marker #71.
XX ds; breast cancer; prognosis; gene expression; diagnosis.
XX Homo sapiens.
XX OS
XX W02004065545-A2.
XX
XX 05-AUG-2004.
XX
XX 15-JAN-2004; 2004WO-US001100.
XX
XX 15-JAN-2003; 2003US-00342887.
XX
XX (ROSE-) ROSETTA INPHARMATICS LLC.
XX (NECA-) NETHERLANDS CANCER INST.
XX
XX

PI Van't Veer LJ, He Y;
XX WPI; 2004-593473/57.
XX
XX Classifying a breast cancer patient according to prognosis comprises
PT determining the similarity between the level of expression of each of
PT five genes in a cell sample taken from patient, to control levels.
XX
XX Disclosure; SEQ ID NO 71; 226pp; English.
XX
XX The invention relates to a method of classifying a breast cancer patient
CC according to prognosis by determining the similarity between the level of
CC expression of each of five genes for which markers are listed in the
CC specification, in a cell sample taken from the breast cancer patient, to
CC control levels of expression for each respective five genes to obtain a
CC patient similarity value. The methods are useful for classifying a breast
CC cancer patient according to prognosis. Kits and computer program products
CC are useful for data analysis using the diagnostic, prognostic and
CC statistical methods of the invention. This sequence corresponds to a
XX marker used in the method of the invention.
SQ Sequence 4955 BP; 1526 A; 803 C; 999 G; 1627 T; 0 U; 0 Other;

Query Match 100.0%; Score 1053; DB 13; Length 4955;
Best Local Similarity 100.0%; Pred. No. 2.1e-290;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCGCACCGGAAAGCGCGGAGCGCGGCTGGATCAGGATCAGATCATCTTCA 60
Db 245 ATGGCGCACCGGAAAGCGCGGAGCGCGGCTGGATCAGGATCAGATCATCTTCA 304

Qy 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCG 120
Db 305 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCG 364

Qy 121 CGGTTTGTCTCTTTTCAATCCAGTACCTCATCTCTCAGATTTTGGAAAAATGATAAAGCAGCAG 180
Db 365 CGGTTTGTCTCTTTTCAATCCAGTACCTCATCTCTCAGATTTTGGAAAAATGATAAAGCAGCAG 424

Qy 181 GCTTCCTTCTGGACAGCAGAGAGTGCATTAATCAAGAGATCTCCCTCAGTGGAAACAG 240
Db 425 GCTTCCTTCTGGACAGCAGAGAGTGCATTAATCAAGAGATCTCCCTCAGTGGAAACAG 484

Qy 241 CTTAAAGCAGATGAGAAAGTACTTCTCATCTCTCAGATTTTGGAGGTTTCAGAGGCT 300
Db 485 CTTAAAGCAGATGAGAAAGTACTTCTCATCTCTCAGATTTTGGAGGTTTCAGAGGCT 544

Qy 301 GGAATTCGTAATGAAAAATTTGGTGGAGCGCTTTAGTTCAGAGGTTTCAGAGGCT 360
Db 545 GGAATTCGTAATGAAAAATTTGGTGGAGCGCTTTAGTTCAGAGGTTTCAGAGGCT 604

Qy 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCAGAAATGTTTCACTCAGAGATGTACAGTTTG 420
Db 605 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCAGAAATGTTTCACTCAGAGATGTACAGTTTG 664

Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTAATGCAATTGAA 480
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTAATGCAATTGAA 724

Qy 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCCCTTGCGATGATAGCAGATAGAAAA 540
Db 725 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCCCTTGCGATGATAGCAGATAGAAAA 784

Qy 541 TCTACTTTTGGGGAAGAGTGGTGGCCCTTGCTGCTGTGAAGAGTGTCTTCTTCAGGA 600
Db 785 TCTACTTTTGGGGAAGAGTGGTGGCCCTTGCTGCTGTGAAGAGTGTCTTCTTCAGGA 844

Qy 601 TCTTTTGTCTGTATTTCTGCTTAAGAAAGAGAGTGTCTTATGCCAGACACCTTTTTC 660
Db 845 TCTTTTGTCTGTATTTCTGCTTAAGAAAGAGAGTGTCTTATGCCAGACACCTTTTTC 904

Qy 661 AATGAACCTCATCAGCAGATGAAGGACTTCACCTGTGACTTTTGTCTGCTGATGTTCCAA 720
Db 905 AATGAACCTCATCAGCAGATGAAGGACTTCACCTGTGACTTTTGTCTGCTGATGTTCCAA 964

Db 905 AATGAATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCTGATGTTCCAA 964
Qy 721 TACTTAGTAATAAGCCCTTCAGAGAAAGGCTCAGGGAGATCATTTGTTGATGCTGTCAAA 780
Db 965 TACTTAGTAATAAGCCCTTCAGAGAAAGGCTCAGGGAGATCATTTGTTGATGCTGTCAAA 1024
Qy 781 ATTGAGCAGGAGTTTAAACAGAGACCTTGCCAGTTGGCCCTCATTTGGAATGAATGCAAT 840
Db 1025 ATTGAGCAGGAGTTTAAACAGAGACCTTGCCAGTTGGCCCTCATTTGGAATGAATGCAAT 1084
Qy 841 TTGATGAAACAGTACATTTGAGTTGTAGCTCAGAGATTACTTGTGGAACCTTGGATTTCTCA 900
Db 1085 TTGATGAAACAGTACATTTGAGTTGTAGCTCAGAGATTACTTGTGGAACCTTGGATTTCTCA 1144
Qy 901 AAGTTTTTTCAGGCAGGAAATCCCTTTTGAATTTATGGAACATTTCTTTAGAGGAA 960
Db 1145 AAGTTTTTTCAGGCAGGAAATCCCTTTTGAATTTATGGAACATTTCTTTAGAGGAA 1204
Qy 961 ACAAAATTTTGTGAGAAACGAGTTTCAGAGATCAGCGTTTGTGAGTTATGCGCAGAAAC 1020
Db 1205 ACAAAATTTTGTGAGAAACGAGTTTCAGAGATCAGCGTTTGTGAGTTATGCGCAGAAAC 1264
Qy 1021 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1053
Db 1265 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1297

RESULT 5

AAF32447
ID AAF32447 standard; cDNA; 1053 BP.

XX AC AAF32447;

XX 18-APR-2001 (first entry)

DE Human ribonucleotide reductase related nucleotide sequence SEQ ID NO:12.

XX Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.

XX Homo sapiens.

XX W0200100799-A1.

XX 04-JAN-2001.

XX 27-JUN-2000; 2000WO-JP004189.

XX 28-JUN-1999; 99JP-00181131.

XX 06-JUL-1999; 99JP-00192391.

XX 21-JAN-2000; 2000JP-00017770.

XX (TAKE) TAKEDA CHEM IND LTD.

XX (NAKA/) NAKAMURA Y.

XX Nakamura Y, Arakawa H, Tanaka H;

XX WPI; 2001-112446/12.

DR Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
XX diagnosis, treatment and prevention of cancer.

XX Claim 6; Page 96-97; 102pp; Japanese.

XX The present invention describes a human ribonucleotide reductase
CC designated TP53R2H. TP53R2H has cytostatic activity. TP53R2H is part of
CC the DNA repair mechanism and its activity is induced by p53. It can be
CC used for the treatment, prevention and diagnosis of a wide range of
CC cancers. The present sequence represents a human ribonucleotide reductase
CC related sequence which is given in the exemplification of the present
XX invention

XX Sequence 1053 BP; 314 A; 191 C; 246 G; 302 T; 0 U; 0 Other;

Query Match 99.8%; Score 1051.4; DB 4; Length 1053;
Best Local Similarity 99.9%; Pred. No. 2.9e-290;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGGGCGACCCGGAAGCGCGGAAGCGCGCGGCTGGATCAGGATCAGAGATCATCTTCA 60
Db 1 ATGGGCGACCCGGAAGCGCGGAAGCGCGCGGCTGGATCAGGATCAGAGATCATCTTCA 60
Qy 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
Qy 121 CGGTTTGTCACTTTTCCAATCCAGTACCTCTGATATTTGGAAATGTATAAACAGGCACAG 180
Db 121 CGGTTTGTCACTTTTCCAATCCAGTACCTCTGATATTTGGAAATGTATAAACAGGCACAG 180
Qy 181 GCTTCTTCTGGACAGCAGAGAGTGCATTTATCAAGGATCTCCCTCATCTGGAAACAAG 240
Db 181 GCTTCTTCTGGACAGCAGAGAGTGCATTTATCAAGGATCTCCCTCATCTGGAAACAAG 240
Qy 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCACTTTAGCCCTTTTTCGAGCCAGTGAT 300
Db 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCACTTTAGCCCTTTTTCGAGCCAGTGAT 300
Qy 301 GGAATTTGTAATGAAATTTTGGTGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
Db 301 GGAATTTGTAATGAAATTTTGGTGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
Qy 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTG 420
Db 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTG 420
Qy 421 CTGATAGACCTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATCAATTTGAA 480
Db 421 CTGATAGACCTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATCAATTTGAA 480
Qy 481 ACCATGCCCTATGTTAAGAAAGGAAAGCAGATTTGGGCTTTGCCGATGATAGCAGATAGAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAGGAAAGCAGATTTGGGCTTTGCCGATGATAGCAGATAGAAA 540
Qy 541 TCTACTTTTGGGAAAGAGTGTGCGCTTTGCTGTGTAAGAGGAGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGTGCGCTTTGCTGTGTAAGAGGAGTTTCTTCTCAGGA 600
Qy 601 TCTTTTGTCTGTATATTTCTGCTAAAGAGAGAGTCTTATGCCAGGACTCATCTTTTCC 660
Db 601 TCTTTTGTCTGTATATTTCTGCTAAAGAGAGAGTCTTATGCCAGGACTCATCTTTTCC 660
Qy 661 AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTGTATGTTCCAA 720
Db 661 AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTGTATGTTCCAA 720
Qy 721 TACTTAGTAATAAGCCCTTCAGAGAAAGGCTTCAGGGAGATCATTTGTTGATGCTGTCAA 780
Db 721 TACTTAGTAATAAGCCCTTCAGAGAAAGGCTTCAGGGAGATCATTTGTTGATGCTGTCAA 780
Qy 781 ATTGAGCAGGAGTTTAAACAGAGACCTTGCCAGTTGGCCCTCATTTGGAATGAATGCAAT 840
Db 781 ATTGAGCAGGAGTTTAAACAGAGACCTTGCCAGTTGGCCCTCATTTGGAATGAATGCAAT 840
Qy 841 TTGATGAAACAGTACATTTGAGTTGTAGCTCAGAGATTACTTGTGGAACCTTGGATTTCTCA 900
Db 841 TTGATGAAACAGTACATTTGAGTTGTAGCTCAGAGATTACTTGTGGAACCTTGGATTTCTCA 900
Qy 901 AAGTTTTTTCAGGCAGGAAATCCCTTTTGAATTTATGGAACATTTCTTTAGAGGAA 960
Db 901 AAGTTTTTTCAGGCAGGAAATCCCTTTTGAATTTATGGAACATTTCTTTAGAGGAA 960
Qy 961 ACAAAATTTTGTGAGAAACGAGTTTCAGAGTATCAGCGTTTGTGAGTTATGCGCAGAAAC 1020
Db 961 ACAAAATTTTGTGAGAAACGAGTTTCAGAGTATCAGCGTTTGTGAGTTATGCGCAGAAAC 1020
Qy 1021 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1053

```
Db 1021 ACAGATAACGCTCTTCACTTGGATGCAGATTTT 1053
|||||
RESULT 6
AAF32440
ID AAF32440 standard; cDNA; 1081 BP.
XX
AC AAF32440;
XX
AT 18-APR-2001 (first entry)
XX
DE Human ribonucleotide reductase related nucleotide sequence SEQ ID NO:4.
XX
KW Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.
XX
OS Homo sapiens.
XX
PN WO200100799-A1.
XX
PD 04-JAN-2001.
XX
PF 27-JUN-2000; 2000WO-JP004189.
XX
PR 28-JUN-1999; 99JP-00181131.
XX
PR 06-JUL-1999; 99JP-00192391.
XX
PR 21-JAN-2000; 2000JP-00017770.
XX
PA (TAKE ) TAKEDA CHEM IND LTD.
XX
PA (NAKA) NAKAMURA Y.
XX
PI Nakamura Y, Arakawa H, Tanaka H;
XX
WPI; 2001-112446/12.
XX
XX
XX Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
PT diagnosis, treatment and prevention of cancer.
XX
XX Example 2; Page 91; 102pp; Japanese.
XX
CC The present invention describes a human ribonucleotide reductase
CC designated TP53R2H. TP53R2H has cytosolic activity. TP53R2H is part of
CC the DNA repair mechanism and its activity is induced by p53. It can be
CC used for the treatment, prevention and diagnosis of a wide range of
CC cancers. The present sequence represents a human ribonucleotide reductase
CC related sequence which is used in an example from the present invention
XX
SQ Sequence 1081 BP; 322 A; 199 C; 252 G; 308 T; 0 U; 0 Other;
Query Match 99.8%; Score 1051.4; DB 4; Length 1081;
Best Local Similarity 99.9%; Pred. No. 2.9e-250;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGGCGACCCGGAAAGCGCGGAGCGCGGCTGGATCAGATGAGAGATCATCTTCA 60
Db 20 ATGGCGACCCGGAAAGCGCGGAGCGCGGCTGGATCAGATGAGAGATCATCTTCA 79
Qy 61 GACACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGTTCGCG 120
Db 80 GACACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGTTCGCG 139
Qy 121 CGGTTTCTCATCTTTCATCCAGTACCTCGATATTTGGAAAATGTATAACAGGCACAG 180
Db 140 CGGTTTCTCATCTTTCATCCAGTACCTCGATATTTGGAAAATGTATAACAGGCACAG 199
Qy 181 GCTTCTCTTCTGGACAGAGAGAGTCTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 240
Db 200 GCTTCTCTTCTGGACAGAGAGAGTTCGACTTATCAAGAGATCTCCCTCACTGGAACAAG 259
Qy 241 CTTAAGCAGATGAGAGTACTTTCATCTCTCACATCTTTAGCCTTTTTTGGAGCCAGTGAT 300
Db 260 CTTAAGCAGATGAGAGTACTTTCATCTCTCACATCTTTAGCCTTTTTTGGAGCCAGTGAT 319
```

```
Qy 301 GGAATTGTAAATGAAAAATTTGGTGGAGCGCTTTAGTCAAGAGGTCAGGTTCCAGAGGCT 360
Db 320 GGAATTGTAAATGAAAAATTTGGTGGAGCGCTTTAGTCAAGAGGTCAGGTTCCAGAGGCT 379
Qy 361 CGCTGTTTCTATGGCTTTCAAATTTCTATCAGAAATGTTTCACTCAGAGATGTACAGTTTG 420
Db 380 CGCTGTTTCTATGGCTTTCAAATTTCTATCAGAAATGTTTCACTCAGAGATGTACAGTTTG 439
Qy 421 CTGATAGACACTTATCATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAAATCAATTTGAA 480
Db 440 CTGATAGACACTTATCATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAAATCAATTTGAA 499
Qy 481 ACCATGCCCTATGTTAAAGAAAAAGCAGATTTGGGCTTTGCCGATGGATAGCAGATAGAAA 540
Db 500 ACCATGCCCTATGTTAAAGAAAAAGCAGATTTGGGCTTTGCCGATGGATAGCAGATAGAAA 559
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAAAGGAGTTTCTTCTCAGGA 600
Db 560 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAAAGGAGTTTCTTCTCAGGA 619
Qy 601 TCTTTTGTCTGCTATATTCTGGCTAAAGAAAGAGAGTCTTTATGCCCAGGACTCACTTTTCC 660
Db 620 TCTTTTGTCTGCTATATTCTGGCTAAAGAAAGAGAGTCTTTATGCCCAGGACTCACTTTTCC 679
Qy 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTCTGCTGATGTTCCAA 720
Db 680 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTCTGCTGATGTTCCAA 739
Qy 721 TACTTAGTAAATAAGCCTTTCAGAAAGAAAGGTCAGGAGAGATCATTTGTGATGCTGTCAA 780
Db 740 TACTTAGTAAATAAGCCTTTCAGAAAGAAAGGTCAGGAGAGATCATTTGTGATGCTGTCAA 799
Qy 781 ATTGAGCAGGAGTTTAAACAGAAAGCCTTGGCAGTTGGCCTCATTTGGAATGAATTCATTT 840
Db 800 ATTGAGCAGGAGTTTAAACAGAAAGCCTTGGCAGTTGGCCTCATTTGGAATGAATTCATTT 859
Qy 841 TTGATGAACAGTACATTTGAGTTTGTAGTCTGACAGATTAATTTGTGGAACTTGGATTTCTCA 900
Db 860 TTGATGAACAGTACATTTGAGTTTGTAGTCTGACAGATTAATTTGTGGAACTTGGATTTCTCA 919
Qy 901 AAGGTTTTTTCAGGAGAAAAATCCTTTGATTTTATGAAAAACATTTCTTTAGAGAAAA 960
Db 920 AAGGTTTTTTCAGGAGAAAAATCCTTTGATTTTATGAAAAACATTTCTTTAGAGAAAA 979
Qy 961 ACAAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGCTTTTGCAGTTATATGGCAGAAACC 1020
Db 980 ACAAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGCTTTTGCAGTTATATGGCAGAAACC 1039
Qy 1021 ACAGATAACGCTCTTCACTTGGATGCAGATTTT 1053
Db 1040 ACAGATAACGCTCTTCACTTGGATGCAGATTTT 1072
RESULT 7
AAH14924
ID AAH14924 standard; cDNA; 1601 BP.
XX
AC AAH14924;
XX
AT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:12810.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
```

PR 29-JUL-1999; 99JP-00248036.
 PR 27-AUG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.
 XX (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Oeuki T;
 XX
 DR WPI; 2001-318749/34.
 XX
 XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.
 XX
 XX Claim 8; SEQ ID NO 12810; 2537pp + Sequence Listing; English.
 XX
 CC The present invention describes primer sets for synthesising 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dr primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesising polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention
 XX
 SQ Sequence 1601 BP; 484 A; 290 C; 348 G; 479 T; 0 U; 0 Other;
 Query Match 99.7%; Score 1049.8; DB 4; Length 1601;
 Best Local Similarity 99.8%; Pred. No. 1e-289;
 Matches 1051; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 361 CGCTGTTTCTATGGCTTTCAAATTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 420
 DB |||||
 QY 400 CGCTGCTTCTATGGCTTTCAAATTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 459
 DB |||||
 QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTTGAA 480
 DB |||||
 QY 460 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTTGAA 519
 DB |||||
 QY 481 ACCATGCCCTATGTTAAGAAAAGGGAATTTTATTTAATGCAATTTGAA 540
 DB |||||
 QY 520 ACCATGCCCTATGTTAAGAAAAGGGAATTTTATTTAATGCAATTTGAA 579
 DB |||||
 QY 541 TCTACTTTTGGGGAAGAGTGTGGCTTTTGTCTGTAGAGAGGAGTTTCTTCTCAGGA 600
 DB |||||
 QY 580 TCTACTTTTGGGGAAGAGTGTGGCTTTTGTCTGTAGAGAGGAGTTTCTTCTCAGGA 639
 DB |||||
 QY 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGTCTTTATGCGCAGGACTCACTTTTTC 660
 DB |||||
 QY 640 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGTCTTTATGCGCAGGACTCACTTTTTC 699
 DB |||||
 QY 661 AATGAACCTCATCAGCAGAGATGAGGACTTCACTGTGACTTTTGTCTGTGATGTTCCAA 720
 DB |||||
 QY 700 AATGAACCTCATCAGCAGAGATGAGGACTTCACTGTGACTTTTGTCTGTGATGTTCCAA 759
 DB |||||
 QY 721 TACTTAGTAAATAAGCCTTTCAGAAAAGGAGTCAAGGAGATCATTTGTTGATGCTGTCAA 780
 DB |||||
 QY 760 TACTTAGTAAATAAGCCTTTCAGAAAAGGAGTCAAGGAGATCATTTGTTGATGCTGTCAA 819
 DB |||||
 QY 781 ATTGAGCAGGAGTTTAAAGAGGAGTTCAGGAGTTCAGGAGTTCAGGAGTTCAGGAGTTC 840
 DB |||||
 QY 820 ATTGAGCAGGAGTTTAAAGAGGAGTTCAGGAGTTCAGGAGTTCAGGAGTTCAGGAGTTC 879
 DB |||||
 QY 841 TTGATGAACAGTACATGAGTTTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGT 900
 DB |||||
 QY 880 TTGATGAACAGTACATGAGTTTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGT 939
 DB |||||
 QY 901 AAGGTTTTTTCAGGAGGAGGAGTTCAGGAGTTCAGGAGTTCAGGAGTTCAGGAGTTCAG 960
 DB |||||
 QY 940 AAGGTTTTTTCAGGAGGAGGAGTTCAGGAGTTCAGGAGTTCAGGAGTTCAGGAGTTCAG 999
 DB |||||
 QY 961 ACAAAATTTCTTTTGAGAAACAGAGTTTTCAGAGTATCAGCGTTTTCAGAGTATGCGCAGAAACC 1020
 DB |||||
 QY 1000 ACAAAATTTCTTTTGAGAAACAGAGTTTTCAGAGTATCAGCGTTTTCAGAGTATGCGCAGAAACC 1059
 DB |||||
 QY 1021 ACAGATAACGTTCTTCACTTGGATGAGTATTT 1053
 DB |||||
 QY 1060 ACAGATAACGTTCTTCACTTGGATGAGTATTT 1092
 DB |||||
 RESULT 8
 ADQ57092
 ID ADQ57092 standard; DNA; 706 BP.
 XX
 AC ADQ57092;
 XX
 DX 21-OCT-2004 (first entry)
 XX
 DE Novel canine microarray-related DNA sequence SeqID8394.
 XX
 KW canine microarray; drug screening; toxicity assay;
 KW environmental pollutant; cellular response; gene expression profile;
 KW toxic response; liver necrosis; fatty liver disease;
 KW protein adduct formation; hepatitis; dog; da.
 XX
 OS Canis familiaris.
 XX
 PN WO2004063324-A2.
 XX
 PD 29-JUL-2004.
 XX
 PF 05-MAY-2003; 2003WO-US013853.
 XX
 PX 03-MAY-2002; 2002US-0377240P.

XX (GENE-) GENE LOGIC INC.
PA (PFIZ) PFIZER PROD INC.
XX
XX Diggins JC, Porter M, Wei T;
XX WPI; 2004-561890/54.
XX
XX New isolated nucleic acid molecule, useful for drug screening and
PT toxicity assays or for assessing the impact, including toxicity, of a
PT compound, pharmaceutical agent or environmental pollutant on a cell or
PT living organism.
XX
XX Claim 1; SEQ ID NO 8394; 41pp; English.
XX
XX This invention is related to a novel isolated canine nucleic acid
CC sequences and the construction of canine microarrays containing a
CC significant portion of the canine genome. The isolated canine nucleic
CC acid sequences of the invention may be useful for drug screening and
CC toxicity assays. The invention is therefore useful for assessing the
CC impact, including toxicity, of a compound, pharmaceutical agent or
CC environmental pollutant on a cell or living organism. The methods are
CC useful for detecting genes that are up- or down-regulated in canines in a
CC disease state. The sequences are useful as diagnostic agents or markers
CC to detect a cellular response in a sample individually or as part of a
CC gene expression profile. It is also useful as a target for agents that
CC modulate gene expression or activity. The database is useful for
CC producing electronic Northernblots that allow the user to determine the cell
CC type or tissue in which a given gene is expressed and to allow
CC determination of the abundance or expression level of a given gene in a
CC particular tissue or cell. The methods are useful for determining the
CC similarity of a toxic response to one or more individual compounds. The
CC methods are useful for predicting at least one toxic response or the
CC likelihood that a compound or test agent will induce various specific
CC pathologies such as those of the liver (liver necrosis, fatty liver
CC disease, protein adduct formation or hepatitis), those of the kidney,
CC heart, brain or testes, or other pathologies associated with at least one
CC of the toxins. The methods are also useful for predicting or elucidating
CC the potential cellular pathways influenced, induced or modulated by the
CC compound or test agent due to the similarity of the expression profile
CC compared to the profile induced by a known toxin. The present sequence is
CC that of a canine DNA sequence which was claimed for use during the
CC production of a canine microarray of the invention.
XX
XX SQ Sequence 706 BP; 208 A; 113 C; 158 G; 227 T; 0 U; 0 Other;

Query Match 61.8%; Score 650.6; DB 13; Length 706;
Best Local Similarity 95.2%; Pred. No. 1.4e-175;
Matches 671; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 349 GTTCCAGAGGCTCGCTTTCTATGCGTTTCAAAATTCATCGAGATGTTCACTCAGAG 408
Db 1 GTTCCAGAGGCTCGCTTTCTATGCGTTTCAAAATTCATCGAGATGTTCACTCAGAG 60
QY 409 ATGTACAGTTGCTGTAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATT 468
Db 61 ATGTACAGTTGCTGTAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATT 120
QY 469 AATCAATTGAACCACTGCTTATGTTAAGAAAGGAAAGGAAATTTTATTGATA 528
Db 121 AATCAATTGAACCACTGCTTATGTTAAGAAAGGAAAGGAAATTTTATTGATA 180
QY 529 GCAGATAGAAATCTACTTTTGGGGAAGAGTGTGGCTTTGCTGTGTAGAGGAGTT 588
Db 181 GAAGATAGAAATCTACTTTTGGGGAAGAGTGTGGCTTTGCTGTGTAGAGGAGTT 240
QY 589 TTCTTCTCAGATCTTTTGGCTGTATATTCGGCTAAAGAGAGAGGTTCTATGCCAGA 648
Db 241 TTCTTCTCAGATCTTTTGGCTGTATATTCGGCTAAAGAGAGAGGTTCTATGCCAGA 300
QY 649 CTCACCTTTTCCAAATGAACTCATCAGCAGATGAAGGACTTCACTGTGACTTTGCTTGC 708
Db 301 CTCACCTTTTCCAAATGAACTCATCAGCAGATGAAGGACTTCACTGTGACTTTGCTTGC 360

QY 709 CTGATGTTCAATACCTTAGTAAATAAGCCTTTCAGAGAAAGGGTCAGGAGATCATTTGTT 768
Db 361 CTGATGTTCAATACCTTAGTAAATAAGCCTTTCAGAGAAAGGGTCAGGAGATCATTTGTT 420
QY 769 GATGCTGTCAAAATTTGAGCAGAGAGTCTTTTAAACAGAGCCTTGGCCCTCATTTGGA 828
Db 421 AATGCTGTGAAATTTGAGCAGAGAGTCTTTTAAACAGAGCCTTGGCCCTCATTTGGA 480
QY 829 ATGAATTCATTTTGTGAAACAGTACATTTAGTTTGTAGCTGCAGAGATTAATCTGTGGA 888
Db 481 ATGAATTCATTTTGTGAAACAGTACATTTAGTTTGTAGCTGCAGAGATTAATCTGTGGA 540
QY 889 CTTCGATTTCTCAAGAGTTTTCAGGAGAAATCTTTTATGATTTTATGAAAAACATTTCT 948
Db 541 CTTCGATTTCTCAAGAGTTTTCAGGAGAAATCTTTTATGATTTTATGAAAAACATTTCT 600
QY 949 TTAGAAGGAAAAACAAATTTCTTTTGAAGAACAGAGTTTCAGAGTATCAGCGTTTTCAGTT 1008
Db 601 TTAGAAGGAAAACTAATTTCTTTTGAAGAACAGAGTTTCAGAGTATCAGCGTTTTCAGTT 660
QY 1009 ATGCGAGAAACACAGATACGTTCTTCACTTGGATGCGAGATTTT 1053
Db 661 ATGCGAGAAACACAGATACGTTCTTCACTTGGATGCGAGATTTT 705
RESULT 9
AAC78111 ID AAC78111 standard; cDNA; 1989 BP.
XX AC AAC78111;
XX DT 08-FEB-2001 (first entry)
XX DE Human cancer associated gene sequence SEQ ID NO:505.
XX Human; cancer associated gene; cancer antigen; detection; cancer;
XX diagnosis; cytostatic; proliferative; vulnerable; immunomodulator;
XX antidiabetic; antiaesthetic; antirheumatic; antiarthritic; antiviral;
XX antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
XX dermatological; neuroprotective; thrombolytic; cosulant; nootropic;
XX vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
XX immune disorder; haematopoietic cell disorder; autoimmune disorder;
XX allergic reaction; graft versus host disease; organ rejection;
XX haemostatic; thrombolytic; cardiovascular disorder; infection;
XX neurological disease; drug screening; ss.
XX Homo sapiens.
XX WO200055350-A1.
XX 21-SEP-2000.
XX 08-MAR-2000; 2000WO-US005882.
XX 12-MAR-1999; 99US-0124270P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Ruben SM;
XX WPI; 2000-587533/55.
XX P-PSDB; AAB43902.
XX Novel isolated nucleic acids comprising sequences encoding peptides
XX useful for treating or diagnosing e.g. cancer.
XX Claim 1; Page 1035-1036; 2352pp; English.
XX AAC77607 to AAC78448 encode the human cancer associated proteins given in
XX AAB43398 to AAB44239. The proteins can have activities based on the
XX tissues and cells the genes are expressed in. Example of activities
XX include: cytostatic; proliferative; vulnerable; immunomodulator;


```
QY 89 ATGAAGACCACTCTCTAAGAAGAGTCTCTCGCGGTTTGTTCATCTTTTCCAAATCCAGTACC 148
Db |||||
QY 283 AGGATGAGCCGCTGCTGAGAGAAACCCCGCGCTTTGTTCATCTTTCCCATCGAGTACC 342
Db |||||
QY 149 CTGATATTTGGAAATGATATAACACGACAGGCTTCCTCTCGACAGCAGAAAGGTGCG 208
Db |||||
QY 343 ATGATATCTGGCAGATGATAGAAGGCGAGAGGCTTCCTTTTGGACCCCGGAGAGGTGG 402
Db |||||
QY 209 ACTTATCAAAAGATCTCCCTCACTGGAACAAGCTTTAAAGCAGATCAGAAAGTACTTCATCT 268
Db |||||
QY 403 ACCTCTCAGAGACATTCAGCACTCGGAATCCCTGAACCCCGAGAGAGATATTTTATAT 462
Db |||||
QY 269 CTCACATCTTAGCCCTTTTTCGAGCCAGTGTGAATTTGTAATCAAAATTTGTTGGAGC 328
Db |||||
QY 463 CCCATGTTCTGGCTTTCTTTTCGACGACGATGGCATAGTAATGAAACTTTGGTGGAGC 522
Db |||||
QY 329 GCTTTAGTCAGAGGTGAGGTTTCAGAGGCTCGCTGTTCTATGCGCTTTCAAATTTCTCA 388
Db |||||
QY 523 GATTTAGCCAAAGAAGTTTCAGATTACAGAAGCCCGCTGTTCTATGGCTTCCAAATTTGCCA 582
Db |||||
QY 389 TCGAAGATGTTTCACTCAGAGATGTACAGTTTGTGCTGATAGACACTTACATCAGATCCCA 448
Db |||||
QY 583 TGGAAACATACATTCGAAATGTATAGTCTCTTATTGACACTTACATAAAGATCCCA 642
Db |||||
QY 449 AGAAAGGGAATTTTATTAAATGCAATTTGAAACCATGCCCTATGTTAAAGAAAAAGCAG 508
Db |||||
QY 643 AAGAAAGGGAATTTCTCTCAATGCCATTGAAACGATGCCCTTGTCTAAGAAGAGGCGAG 702
Db |||||
QY 509 ATTGGGCTTGGGATGAGATGAGAGATGAAATCTACTTTTGGGAAAGAGTGTGGGCT 568
Db |||||
QY 703 ACTGGGCTTGGGCTGATGGGACAAAGAGGCTACCTATGGTGAACGTGTTGTAGGCT 762
Db |||||
QY 569 TTGCTGCTGTAGAGGAGTTTCTTCTCAGATCTTTTGTGCTGATATTTCTGGCTTAAGA 628
Db |||||
QY 763 TTGCTGCTGTAGAGGAGTTTCTTCTTCCGGTCTTTTTCGCTGATATTTCTGGCTCAAGA 822
Db |||||
QY 629 AGAGAGGTCTTATGCGAGGACTCACTTTTCCAAATGAACTCATCAGCAGAGATGAAGAC 688
Db |||||
QY 823 AACGAGGACTGATGCTGGCTCACAATTTCTAATGAATTTATAGCAGATGAGGTT 882
Db |||||
QY 689 TTCACTGTGACTTTGCTTGTGCTGATGTTCCAAATCTTAGTAAATAAGCCTTTCAGAAGAA 748
Db |||||
QY 883 TACACTGTGATTTTGTCTGCTGATGTTTCAAAACACCTGTTACACAAACCATCGGAGGAGA 942
Db |||||
QY 749 GGGTCAGGAGATCATTTGCTGATGCTGTCAAAATGACAGAGTTCCTTACAGAGGCT 808
Db |||||
QY 943 GAGTAAGAAGAAATAATATCAATGCTGTTCCGATAGAAACAGGAGTTCCTCACTGAGGCT 1002
Db |||||
QY 809 TGCCAGTTGGCTCATTTGGAATGAATTCATTTTGAATGAACAGTACATTCAGTTCAGTTGTAG 868
Db |||||
QY 1003 TGCCGTGAAGCTCATTTGGGATGAATTCGACTCTAATGACCAATACATTCAGTTCAGTTGTAG 1062
Db |||||
QY 869 CTGACAGATTAATTTGTGGAATCTTGGAATTTCTCAAAAGTTTTCAGGCGAGAAATCTTTTG 928
Db |||||
QY 1063 CAGACAGACTTATGCTGGAATCGGCTTTTAGCAAGTTTTCAGAGTACAGAACCCATTTG 1122
Db |||||
QY 929 ATTTTATGGAACAAATTTCTTAGAAGAAACAAATTTCTTTGAGAAACAGTTTCAG 988
Db |||||
QY 1123 ACTTTATGGAGATTAATTTCTACTGGAAGGAAAGACTTAACCTCTTTTGAAGAGAGTAGGCG 1182
Db |||||
QY 989 AGTATCAGGCTTTTTCAGTTATGGCAGAAACACACAGATAAGCTCTTACCTTGTGATGAG 1048
Db |||||
QY 1183 AGTATCAGAGATGGGAGTATGTCAGTCCAAACAGAGATTTCTTTTACCTTGTGATGCTG 1242
Db |||||
QY 1049 ATTT 1052
Db |||||
QY 1243 ACTT 1246
Db |||||
```

RESULT 11
ADJ56536
ID ADJ56536 standard; cDNA; 2216 BP.
XX

```
AC ADJ56536;
XX
DT 06-MAY-2004 (first entry)
XX
DE Murine cDNA differentially expressed in MYCN activated cells SeqID 342.
XX
KW mouse; murine; differential expression; transactivator; proto-oncogene;
XX neuroblastoma; small cell lung cancer; cytostatic; gene therapy; ss;
XX MYCN activated cell.
XX
OS Mus musculus.
XX
PN US2003119009-A1.
XX
PD 26-JUN-2003.
XX
PF 25-FEB-2002; 2002US-00084817.
XX
PR 23-FEB-2001; 2001US-0270784P.
XX
PA (STUA/) STUART S G.
PA (NUCH/) NUCHTERN J G.
PA (PLON/) PLON S E.
PA (SHOH/) SHOHET J M.
XX
XX Stuart SG, Nuchtern JG, Plon SE, Shohet JM;
XX WPI; 2003-635698/60.
XX
XX New genes regulated by MYCN activation, useful in gene therapy,
XX particularly for treating a subject with e.g. neuroblastoma or other
XX cancers, or for diagnosing, staging or monitoring the treatment of the
XX cancer.
XX
FS Claim 1; SEQ ID NO 342; 27pp; English.
XX
XX This invention relates to novel isolated cDNAs that are differentially
XX expressed in MYCN activated cells. Specifically, it refers to
XX polynucleotide sequences that exhibit differential expression patterns in
XX cells activated by the transactivator MYCN, where MYCN is a proto-
XX oncogene that is amplified in neuroblastoma cells and is common in small
XX cell lung cancers. The present invention describes these cDNA molecules
XX as useful for in hybridisation assays to detect expression of nucleic
XX acids (or complementary nucleic acids) in a present in a given sample, as
XX well as for screening assays by identifying molecules or compounds that
XX specifically bind the cDNA as a ligand and modulate function or activity.
XX Accordingly, these compositions exhibit cytostatic activity and can also
XX be used for gene therapy purposes. This polynucleotide sequence is a cDNA
XX that is differentially expressed in MYCN activated cells, given in an
XX exemplification of the invention. NOTE: This sequence does not appear in
XX the printed specification but has been obtained in electronic format from
XX the US Patent Office at
XX ftp.segdata.uspto.gov/sequence.html?DocID=20030119009.
XX
SQ Sequence 2216 BP; 593 A; 476 C; 524 G; 623 T; 0 U; 0 Other;
```

```
Query Match 54.9%; Score 578.4; DB 10; Length 2216;
Best Local Similarity 75.0%; Pred. No. 1.2e-154;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

QY 89 ATGAAGACCACTCTCTAAGAAGAGTCTCTCGCGGTTTGTTCATCTTTCCAAATCCAGTACC 148
Db |||||
QY 284 AGGATGAGCCGCTGCTGAGAGAAACCCCGCGCTTTGTTCATCTTTCCCATCGAGTACC 343
Db |||||
QY 149 CTGATATTTGGAAATGATATAACACGACAGGCTTCCTCTCGACAGCAGAAAGGTGCG 208
Db |||||
QY 344 ATGATATCTGGCAGATGATAGAAGGCGAGAGGCTTCCTTTTGGACCCCGGAGAGGTGG 403
Db |||||
QY 209 ACTTATCAAAAGATCTCCCTCACTGGAACAAGCTTTAAAGCAGATCAGAAAGTACTTCATCT 268
Db |||||
QY 404 ACCTCTCAGAGACATTCAGCACTCGGAATCCCTGAACCCCGAGAGAGATATTTTATAT 463
Db |||||
QY 269 CTCACATCTTAGCCCTTTTTCGAGCCAGTGTGAATTTGTAATCAAAATTTGGTGGAGC 328
Db |||||
```

Db 464 CCCAGTCTGGCTTCTTTGAGCAAGCGATGGCATAGTAATGAAACCTTGGTGGAGC 523
Qy 329 GCTTTAGTCAGAGGTGCGAGTTCAGAGGCTCGCTGTTCTTATGGCTTTCAAATTTCTCA 388
Db 524 GATTTAGCAAGAAGTTCAATTCAGAAAGCCGCTGTTCTATGGCTTCCAAATTCGCA 583
Qy 389 TCGAGAAATGTTCACTCAGAGATGTACAGTTTGTCTGATAGACACATTTACATCAGAGATCCCA 448
Db 584 TGGAAACATACATCTCGAAATGTATAGTCTTCTTTATTTGACACTTACATAAAGATCCCA 643
Qy 449 AGAAAGGGAATTTTATTTAATGCAATGAAACCATGCCCCATGTTTAAAGAAAAAGCAG 508
Db 644 AAGAAAGGGAATTTCTCTTCAATGCAATTTGAAACGATGCTTGTCTCAAGAAAGAGGCGAG 703
Qy 509 ATTGGGCTTCGCGATGATACAGATACAAATCTACTTTTGGGGAAGAGTGGTGGCT 568
Db 704 ACTGGGCTTCGCGATGATGCGGCAAGAGGCTTACCTATGGTGAACGTGTGTAGGCT 763
Qy 569 TTGCTGCTGTAGAAGGAGTCTTCTTCTCAGGATCTTTTGGCTATATTTCTGGCTAAAGA 628
Db 764 TTGCTGAGTGAAGGCAATTTCTTTCCGGTCTTTTGGCTGATATTTCTGGCTCAGA 823
Qy 629 AGAGAGTCTTATGCCAGGACTCACTTTTCCAAATGAACTCATCAGCAGAGATGAAGCAC 688
Db 824 AACGAGGACTGATGCTGGCTCACAATTTCTAATGAACTTATTAGCAGAGATGAGGCTT 883
Qy 689 TTCACTGACTTTGCTGCTGATGTCCAAATCTAGTAAATTAAGCTTCAAGAGAA 748
Db 884 TACACTGTGATTTTGGCTGCTGATGTTCAAACACCTGGTACACAAACCATCGGAGGAGA 943
Qy 749 GGGTCAGGAGATCAATTTGATGCTGTCAAAATTTGAGCAGAGTTTTAAACAGAGCT 808
Db 944 GAGTAAGAGAAATTAATCAATGCTGTTCGATAGAACAGAGTCTCTACTGAGGCT 1003
Qy 809 TGCCAGTGGCTCATTTGGAATGAATTCATTTTGAATGAACAGTACATTCAGTTTGTAG 868
Db 1004 TGCCTGTGAAGCTCATTTGGGATGAATTCGACTTAATGAAGCAATACATTCAGTTTGTG 1063
Qy 869 CTGACAGATTAATTTGGAAGTGGATGCTCAAGGTTTTCAGGCGAGAAATCTTTTG 928
Db 1064 CAGACAGACTTATGCTGGAAGTGGGTTTTCAGAAAGTTTTCAGAGTGAAGAACCAATTTG 1123
Qy 929 ATTTTATGGAACATTTCTTGAAGGAAACAAATTTCTTGAAGAACGCTTTCAG 988
Db 1124 ACTTATGGAATTAATTTCTGGAAGAAAGACTAACTCTTTTGAAGAGAGTGAAGCG 1183
Qy 989 AGTATCAGCGTTTTCAGTTATGTCAGAGAAACCAAGATAAGCTTTCACCTTGCATGCAG 1048
Db 1184 AGTATCAGAGATGGGAGTGAATGTCAGTCCACAGAGAAATTTCTTTACCTTGGATGCTG 1243
Qy 1049 ATTT 1052
Db 1244 ACTT 1247

RESULT 12

AAS44917/C

ID AAS44917 standard; DNA; 2482 BP.

XX AC AAS44917;

XX AC AAS44917;

DT 18-DEC-2001 (first entry)

XX Human contig polynucleotide sequence #170.

XX Mammal; human; rhesus monkey; baker's yeast; fission yeast; Norway rat;
XX mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;
XX cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;
XX nervous system disorder; inflammatory disorder; cell differentiation; ds;
XX angiogenesis; stem cell growth factor; activin; inhibin; cartilage; burn;
XX genetic disorder; bone regeneration; tendon; ligament; tissue repair;
XX cytosolic; antirheumatic; antiarthritic; vulnery; antiinflammatory;

KW antibacterial; immunosuppressive; vasotropic; antiparkinsonian;
KW neuroprotective; osteopathic; antidiabetic; antiasthmatic; antiallergic;
KW immunostimulant; analgesic; gene therapy.

OS Homo sapiens.

XX Synthetic.

XX WO200164834-A2.

XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US004926.

XX 28-FEB-2000; 2000US-00515126.

XX 18-MAY-2000; 2000US-00577409.

XX 17-JUN-2000; 2000US-00597707.

XX 14-JUL-2000; 2000US-00616807.

XX 19-SEP-2000; 2000US-00664641.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;

PI Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;

PI Drmanac R;

XX WPI; 2001-589862/66.

XX P-PSDB; AAU28017.

XX Novel polypeptides and nucleic acids obtained from cDNA libraries
XX prepared from various human tissues, for diagnosis, treatment of cancer,
XX neurological, inflammatory disorders and for use in arrays for detection.

XX Claim 1; SEQ ID NO 514; 153pp; English.

XX Sequences AAS44576-AAS44919 represent full-length polynucleotides and
XX contig polynucleotides encoding polypeptides of the invention. The DNA
XX and protein sequences are useful for the treatment, diagnosis and
XX prevention of various types of disorder in a mammalian subject such as a
XX human, dog, monkey, mouse, hamster or rat. The disorders include cancers
XX such as leukaemia, lymphoma and neuroblastoma, autoimmune disorders such
XX as multiple sclerosis, connective tissue disease, rheumatoid arthritis,
XX diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system
XX disorders such as Parkinson's disease, Alzheimer's disease, Huntington's
XX chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and
XX Wernicke disease, inflammatory disorders such as nephritis, Crohn's
XX disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory
XX bowel disease. The sequences exhibit activity relating to angiogenesis,
XX cell proliferation, cell differentiation, stem cell growth factor,
XX activin or inhibin. Therefore, they can be used to manipulate stem cells
XX in culture to give rise to neuroepithelial cells that can be used to
XX augment or replace cells damaged by illness, accidental damage or genetic
XX disorders. The sequences may also be used for regeneration of bone,
XX cartilage, tendons and ligaments and in tissue repair and burn healing.
XX Note: Some sequences for this patent did not form part of the printed
XX specification, but were obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 2482 BP; 707 A; 598 C; 522 G; 655 T; 0 U; 0 Other;

XX Query Match 54.9%; Score 578.4; DB 4; Length 2482;

XX Best Local Similarity 75.0%; Pred. No. 1.2e-154;

XX Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

Qy 89 ATGAAGAGCCACTCTCTAAGAAAGAGTCTTCGCGGTTTGTCTATCTTTTCCAAATCCAGTACC 148

Db 2086 AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTGTCTATCTTTCCCATCGAGTACC 2027

Qy 149 CTGATATTGGAATAATGATATAAAGGCAAGGCTTCTCTTGGACAGAGAGAGTGG 208

Db 2026 ATGATATCTGCAGATGATATAAGAGGCGAGAGCTTCTCTTTGGACCGCGAGAGTGG 1967

Qy 209 ACTTATCAAGAGATCTCCCTCAGTGAACAGCTTAAGACAGATGAGAGTACTTCATCT 268

Db 1966 ACCTCTCCAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT 1907
Qy 269 CTACATCTTTAGCCTTTTTCAGCCAGTGAATTTGTAATGAAATTTGGTGGAGC 328
Db 1906 CCATGTTCTGGCTTTCTTTCAGACAGGAGTGCATAGTAATGAAACTTTGGTGAGC 1847
Qy 329 GCTTTAGTCAGGAGTTCAGGTTCCAGAGGCTCGCTGTTCTATGGCTTTCAAAATTTCTCA 388
Db 1846 GATTTAGCCAGAGTTCAGATTACAGAACCCGCTGTTCTATGGCTTCCAAATTTGCCA 1787
Qy 389 TCGGAATGTTCACTCAGAGATGATACAGTTTGTGCTGATAGACACTTACATCAGAGATCCCA 448
Db 1786 TGGAAAAACATACATTTCTGAATGTATAGTCTTCTTATTTGACACTTACATAAAGATCCCA 1727
Qy 449 AGAAAAGGGAATTTTATTTAATCAATTCGAATCCATGCAACCATGCGCTATGTTAAGAAAAGACAG 508
Db 1726 AAGAAAGGAATTTCTCTTCAATGCCATTTGAACGATGCGCTTGTGTCAGAAAGAGGAGCAG 1667
Qy 509 ATTGGGCGCTTTCGATAGGATAGCAGATAGAAAATCTACTTTTGGGGAAGAGTGGTGGCCT 568
Db 1666 ACTGGGCTTTCGCTGGATTCGGGACAAAGAGGCTACTATGTTGAACGTTGTAGCCT 1607
Qy 569 TTGCTGCTGATAGAGGAGTTTCTTCTCAGAGTCTTTTGTGCTATATTTCTGCTAAAGA 628
Db 1606 TTGCTGAGTGAAGGCAATTTCTTTTCGGGTTCTTTTGGCTCGATATTTCTGGCTCAAGA 1547
Qy 629 AGAGAGTCTTATGCCAGGACTCACTTTTCCATGAACTCATCAGCAGAGATCAAGGAC 688
Db 1546 AACGAGGACTGATGCGCTGCATTTCTAATGAATCTTATTAGCAGAGATGAGGTT 1487
Qy 689 TTCACTGTGACTTTGCTTGGCTGATGTTTCAATATCTTAGTAAATAAGCCTTTCAGAAAGA 748
Db 1486 TACACTGTGATTTTGTCTGCTGATGTTTCAACACCTGTTACACAAACCATCGAGGAGA 1427
Qy 749 GGGTCAGGAGATCATTTGTCATGCTGTCAAAATGAGCAGAGTTTTAAACAGAGCCT 808
Db 1426 GAGTAAGAGAAATAATATCAATGCTGTTCCGATAGAAACAGGAGTTCCTACTGAGGCT 1367
Qy 809 TGCAGTGGCCTCATTTGGATGAATGTCATTTTGTATGAACACATACATTCAGTTTCTAG 868
Db 1366 TGCCTGTGAAGCTCATTTGGGATGAATGTCATCTTAATGAAGCAATACATTCAGTTTCTGG 1307
Qy 869 CTGACAGATTAATTTGGAATCTGATTTCTCAAAAGTTTTTCAGGCGAGAAATCCTTTTG 928
Db 1306 CAGACAGACTTATGCTGGAATCTGGTTTACCAAGTTTTTCAGAGTAGAACCCATTTG 1247
Qy 929 ATTTTATGGAATAATTTCTTTTGAAGAAACAAATTTCTTTGAGAAACGAGTTTTCAG 988
Db 1246 ACTTTATGGAGATATTTTCACTGGAAGGAAAGACTAACTTCTTTGAGAGAGAGTAGGCG 1187
Qy 989 AGTATCAGCGTTTTCAGTTTATGCGAGAAACCAACAGATACGTTCTTACCTTGGATGCG 1048
Db 1186 AGTATCAGAGATGGGAGTGTGTCAGTCCACAGAGAAATTTCTTTTACCTTGGATGCTG 1127
Qy 1049 ATTT 1052
Db 1126 ACTT 1123

RESULT 13

ABL65414

ID ABL65414 standard; DNA; 2500 BP.

XX AC ABL65414;

XX DT 15-MAY-2002 (first entry)

XX DE Lung cancer related gene sequence SEQ ID NO:3751.

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.

XX Homo sapiens.
OS WO200194629-A2.
PN 13-DEC-2001.
PD 30-MAY-2001; 2001WO-US010838.
XX 05-JUN-2000; 2000US-0209473P.
XX 05-JUN-2000; 2000US-0209531P.
XX 18-SEP-2000; 2000US-0233133P.
XX 18-SEP-2000; 2000US-0233617P.
XX 20-SEP-2000; 2000US-0234009P.
XX 20-SEP-2000; 2000US-0234034P.
XX 20-SEP-2000; 2000US-0234052P.
XX 22-SEP-2000; 2000US-0234509P.
XX 22-SEP-2000; 2000US-0234567P.
XX 25-SEP-2000; 2000US-0234923P.
XX 25-SEP-2000; 2000US-0234924P.
XX 25-SEP-2000; 2000US-0235077P.
XX 25-SEP-2000; 2000US-0235082P.
XX 25-SEP-2000; 2000US-0235134P.
XX 25-SEP-2000; 2000US-0235280P.
XX 26-SEP-2000; 2000US-0235637P.
XX 26-SEP-2000; 2000US-0235638P.
XX 27-SEP-2000; 2000US-0235711P.
XX 27-SEP-2000; 2000US-0235720P.
XX 27-SEP-2000; 2000US-0235840P.
XX 27-SEP-2000; 2000US-0235863P.
XX 28-SEP-2000; 2000US-0236032P.
XX 28-SEP-2000; 2000US-0236033P.
XX 28-SEP-2000; 2000US-0236034P.
XX 28-SEP-2000; 2000US-0236109P.
XX 28-SEP-2000; 2000US-0236111P.
XX 29-SEP-2000; 2000US-0236842P.
XX 29-SEP-2000; 2000US-0236891P.
XX 02-OCT-2000; 2000US-0237172P.
XX 02-OCT-2000; 2000US-0237173P.
XX 02-OCT-2000; 2000US-0237278P.
XX 02-OCT-2000; 2000US-0237294P.
XX 02-OCT-2000; 2000US-0237295P.
XX 02-OCT-2000; 2000US-0237316P.
XX 03-OCT-2000; 2000US-0237425P.
XX 03-OCT-2000; 2000US-0237598P.
XX 03-OCT-2000; 2000US-0237604P.
XX 03-OCT-2000; 2000US-0237606P.
XX 03-OCT-2000; 2000US-0237608P.
XX 01-NOV-2000; 2000US-0244867P.
XX 01-NOV-2000; 2000US-0245084P.
XX (AVAL-) AVALON PHARM.
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX WPI; 2002-188264/24.
XX Screening for anti-neoplastic agent involves exposing cells to a chemical
XX agent to be tested for anti-neoplastic activity, and determining a change
XX in expression of a gene of a signature gene set.
XX Claim 1; SEQ ID NO 3751; 44pp; English.
XX The present invention describes a method (M1) for screening for an anti-
XX neoplastic agent. The method involves exposing cells to a chemical agent
XX to be tested for anti-neoplastic activity, determining a change in
XX expression of at least one gene (I) of a signature gene set, where (I)
XX comprises a sequence (S) selected from 8447 sequences (given in ABL61664
XX to ABL70110), or is at least 95% identical to (S), where a change in
XX expression is indicative of anti-neoplastic activity. (I) has cytostatic
XX activity and can be used in gene therapy. M1 can be used for screening an

CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of MI, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. MI can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC tumour
XX
SQ Sequence 2500 BP; 679 A; 522 C; 595 G; 704 T; 0 U; 0 Other;

Query Match 54.9%; Score 578.4; DB 6; Length 2500;
Best Local Similarity 75.0%; Pred. No. 1.2e-154;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

Qy	89	ATGAGAGCCACTCTCTAGAAAGAGTTCTCCCGGTTTGTGATCTTTCCATCCAGTACC	148
Db	397	AGGATGAGCCGCTGCTGAGAGAAACCCCGCCGCTTTGTATCTTCCCATCGAGTACC	456
Qy	149	CTGATATTGGAAATGATATAACAGGCACAGGCTTCTCTGGACAGCAGAGAGTGC	208
Db	457	ATGATATCTGCAGATGATAGAGAGCAGAGGCTTCTTTGGACCCCGAGAGGTG	516
Qy	209	ACTTATCAAGAGATCTCCCTCACTGGAAACAAGCTTAAAGCAGATGAGAGTACTTCATCT	268
Db	517	ACCTCTCAAGGACATTCAGCACTGGGAATCCCTGNAACCCGAGGAGATATTTATAT	576
Qy	269	CTCAGATCTTAGCCCTTTTTCAGCCAGTATGGAATGTAATGAAATTTGGTGAGC	328
Db	577	CCCATGTTCTGGCTTTCTTTGACGACGATGGCATAGTAAATGAAATCTGGTGAGC	636
Qy	329	GCTTTAGTCAGAGTGAGTTCAGAGGCTCGCTGTTTCTATGGCTTTCAAAATCTCA	388
Db	637	GATTTAGCCAAAGAGTTCAGATTAAGAGCCCGCTGTTCTATGGCTTCCAAATGCGCA	696
Qy	389	TCGAGATGTTCACTCAGAGATGTACAGTTTGTGTATAGACACTTTACATCAGAGATCCCA	448
Db	697	TGGAACACATACATCTGAAATGTATAGTCTTCTTATGACACTTACATAAAGATCCCA	756
Qy	449	AGAAAGGGAATTTTATTAATGCAATGAAACATGCCCCTATGTTAAGAAAAAAGCAG	508
Db	757	AAGAAAGGGAATTTCTCTCAATGCCATTTGAAACGATGCTTGTCTCAAGAAAGGCAG	816
Qy	509	ATTGGGCTTCGATGATAGACATAGAAATCTACTTTTGGGGAAGAGTGTGGCCT	568
Db	817	ACTGGGCTTCGCTGATGATGGGACAAAGAGGCTACCTATGGTGAAGCTTTGTAGCCT	876
Qy	569	TTGCTGCTGTAGAAGGAGTTTCTTCTCAGGATCTTTTGTGCTATATTTCTGGCTAAAGA	628
Db	877	TTGCTGAGTGAAGGCAATTTCTTTTCCGGTCTTTTGGCTCGATATTTCTGGCTCAGA	936
Qy	629	AGAGAGTCTTATGCCAGGACTCACTTTTTCATGAACTCATCAGCAGAGATGAAGGAC	688
Db	937	AACGAGGACTGATGCTGGCTCACAATTTCTAATGAATTTATTAGCAGATGAGGGTT	996
Qy	689	TTCACTGTGATTTGCTGCTGATGTTCCATGTTAGTAAATAGCTTCAGAGAA	748
Db	997	TACACTGTGATTTTGGCTTGGCTGATGTTTCAACACCTGGTACACAAACCATCGGAGGA	1056
Qy	749	GGGTGAGGAGATCATTTGTTGATGCTGTCATAAATTTGAGCAGGAGTTTAAACAGCCT	808
Db	1057	GAGTAAGAGAAATTAATTAATGCTGTTGGAATAGAACAGGATTCCTACTGAGGCT	1116
Qy	809	TGCCAGTTGGCTCATTTGGAATGAATTCATTTTGTATGAACAGTACATTCAGTTTGTAG	868
Db	1117	TGCCTGTGAGACTCATTTGGGATGAATTTGCACTCTAATGAAGCAATACATTGAGTTTGG	1176
Qy	869	CTGACAGATTAATTTGTTGGAATCTCAAGGTTTTCAGGCGAGAAATCTTTTG	928
Db	1177	CAGACAGACTTATGCTGGAATCTGGGTTTATGCAAGGTTTTCAGAGTAGAGAACCAATTG	1236
Qy	929	ATTTTATGGAACATTTCTTTAGAGGAAAAACAAATTTCTTTGAGAAACGAGTTTCAG	988

Db	1237	ACTTTATGGAATATTTTCACTGGAGGAAAGACTAACTTCTTTGAGAGAGTAGGCG	1296
Qy	989	AGTATCAGCGTTTTCAGTTATGCGAGAAACACACAGATAACGTTTTCACCTTCGATGCG	1048
Db	1297	AGTATCAGAGGATGGAGTGATGTCAAGTCCACACAGAGAAATTTTACCTTCGATGCTG	1356
Qy	1049	ATTT 1052	
Db	1357	ACTT 1360	
RESULT 14			
ABL66517			
ID	ABL66517 standard; DNA; 2500 BP.		
XX	AC ABL66517;		
XX	XX		
DT	15-MAY-2002 (first entry)		
XX	Lung cancer related gene sequence SEQ ID NO:4854.		
DE	Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;		
KW	stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;		
KW	cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;		
KW	gene; da.		
XX	Homo sapiens.		
OS	Homo sapiens.		
XX	WO200194629-A2.		
PN	13-DEC-2001.		
PD	30-MAY-2001; 2001WO-US010838.		
XX	05-JUN-2000; 2000US-0209473P.		
PR	05-JUN-2000; 2000US-0209531P.		
PR	18-SEP-2000; 2000US-0233133P.		
PR	18-SEP-2000; 2000US-0233617P.		
PR	20-SEP-2000; 2000US-0234009P.		
PR	20-SEP-2000; 2000US-0234034P.		
PR	22-SEP-2000; 2000US-0234509P.		
PR	22-SEP-2000; 2000US-0234567P.		
PR	25-SEP-2000; 2000US-0234923P.		
PR	25-SEP-2000; 2000US-0234924P.		
PR	25-SEP-2000; 2000US-0235077P.		
PR	25-SEP-2000; 2000US-0235082P.		
PR	25-SEP-2000; 2000US-0235134P.		
PR	25-SEP-2000; 2000US-0235280P.		
PR	26-SEP-2000; 2000US-0235637P.		
PR	26-SEP-2000; 2000US-0235638P.		
PR	27-SEP-2000; 2000US-0235711P.		
PR	27-SEP-2000; 2000US-0235720P.		
PR	27-SEP-2000; 2000US-0235840P.		
PR	27-SEP-2000; 2000US-0235863P.		
PR	28-SEP-2000; 2000US-0236028P.		
PR	28-SEP-2000; 2000US-0236032P.		
PR	28-SEP-2000; 2000US-0236033P.		
PR	28-SEP-2000; 2000US-0236034P.		
PR	28-SEP-2000; 2000US-0236109P.		
PR	28-SEP-2000; 2000US-0236111P.		
PR	29-SEP-2000; 2000US-0236842P.		
PR	29-SEP-2000; 2000US-0236891P.		
PR	02-OCT-2000; 2000US-0237172P.		
PR	02-OCT-2000; 2000US-0237173P.		
PR	02-OCT-2000; 2000US-0237278P.		
PR	02-OCT-2000; 2000US-0237294P.		
PR	02-OCT-2000; 2000US-0237295P.		
PR	02-OCT-2000; 2000US-0237316P.		
PR	03-OCT-2000; 2000US-0237425P.		
PR	03-OCT-2000; 2000US-0237598P.		
PR	03-OCT-2000; 2000US-0237604P.		

PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
XX

(AVAL-) AVALON PHARM.

XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
XX Soppet DR, Weaver Z;

XX WPI; 2002-188264/24.

PT Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.

XX Claim 1; SEQ ID NO 4196; 44pp; English.

XX The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC tumour

XX Sequence 2500 BP; 679 A; 522 C; 595 G; 704 T; 0 U; 0 Other;

Query Match 54.9%; Score 578.4; DB 6; Length 2500;
Best Local Similarity 75.0%; Pred. No. 1.2e-154;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

Qy 89 ATGAAGACCCACTCTTAAGAAAGAGTTCTCGCCGGTTTGTCTCTTTCCATCCAGTACC 148
Db |||||
397 AGGATGAGCCGCTGCTGAGAGAAACCCCGCGCTTGTCTCTCTCCCATCGAGTACC 456
Qy 149 CTGATATTGGAAATGTATTAACAGGCACAGGCTTCTCTGGACAGCAGAGAGTCG 208

Db |||||
457 ATGATATCTGGCAGATGTATTAAGAAAGCAGAGGCTTCTCTTTGGACCGCCGAGGAGTTG 516
Qy |||||
209 ACTTATCAAGGATCTCCCTCACTGGAAACAGCTTTAAAGCAGATCAGAGTACTTTCATCT 268
Db |||||
517 ACCTCTCAAGGACANTCAGCACCTGGGAATCCCTCAAAACCCGAGGAGATATTTTATAT 576
Qy |||||
269 CTCACATCTTAGCCCTTTTTCGAGCCAGTGTAAATGTAATGAAATTTTGGTGAGC 328
Db |||||
577 CCCATGTTCTGGCTTCTTTGACGACAGGATGGCATAGTAATGAAATTTTGGTGAGC 636
Qy |||||
329 GCTTTAGTCAGGAGGTGAGGTTCCAGAGGCTCGCTGTTTCTATGCGCTTTCAAAATCTCA 388
Db |||||
637 GATTTAGCCAAGAGTTCAGATTACAGAAAGCCCGCTGTTCTATGGCTTCCAAATTTGCCA 696
Qy |||||
389 TCGAGAAATGTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTTACATCAGAGATCCA 448
Db |||||
697 TGGAAACATACATTTCTGAAATGTATAGTCTTCTTATTTGACACTTTACATAAAGATCCA 756
Qy |||||
449 AGAAAGGGAATTTTATTTAATGCAATTTGAAACCATGCCCTATGTTTGGGAAAGAGTGGTGCCT 508
Db |||||
757 AAGAAAGGGAATTTCTCTTCAATGCCATTGAAACGATGCCCTTGTCTAAGAAAGAGGAC 816
Qy |||||
509 ATTGGGCTTTCGAGTAGATAGACATAGAAATCTACTTTTGGGAAAGAGTGGTGCCT 568
Db |||||
817 ACTGGGCTTTCGAGTAGATAGACATAGAAAGGCTTACCTATGTTGAAAGCTGTTAGCCT 876
Qy |||||
569 TTGCTGCTGTAGAAAGAGTCTTCTTCTCAGAGATCTTTTGTCTGTATATTTCTGGCTTAAAGA 628
Db |||||
877 TTGCTGAGTGGAAAGGCAATTTCTTTTCCGGTCTCTTTTCCGGTCGATATTTCTGGCTCAAGA 936
Qy |||||
629 AGAGAGTCTTATGCCAGGACTCACTTTTCCAAATGAACTCATCAGCAGAGATGAAGAC 688
Db |||||
937 AACGAGGACTGATGCCCTGCCCTCACAATTTCTAATGAACCTTATTAGCAGAGATGAGGTT 996
Qy |||||
689 TTCACCTGTGACTTTGCTTGCCTGATGTTCCAAATCTTAGTAAATAAGCCTTCAGAAAGAA 748
Db |||||
997 TACACTGTGATTTGCTTGCCTGATGTTCAAAACCTGTTTACAAACATCGGAGGAGA 1056
Qy |||||
749 GGGTCAGGAGATCATTTGTGATGCTGTCAAAATTTGAGCAGGAGTTTAAACAGAGCCT 808
Db |||||
1057 GAGTAAGAGAAATAATTAATCAATGCTGTTCCGATAGAAACAGGAGTTCTCTCACTGAGGCT 1116
Qy |||||
809 TGCAGTTGGCTCATTTGGAAATGATTCATTTTGTATGAAACAGTACATTTGAGTTTGTAG 868
Db |||||
1117 TGCCTGTGAAGCTCATTTGGGATGAAATGCACTCTAATGAAGCAATACATTTGAGTTTGTGG 1176
Qy |||||
869 CTGACAGATTACTTGTGGAACCTGCAATTTCTCAAAGTTTTCAGGCAGCAAAATCCTTTTG 928
Db |||||
1177 CAGACAGATTATGCTGGAACCTGGGTTTTCAGCAAGTTTTCAGAGTAGAGAACCCATTTG 1236
Qy |||||
929 ATTTTATGGAACATTTCTTTTGAAGGAAACAAATTTCTTTTGAAGAAACAGATTTCAG 988
Db |||||
1237 ACTTTATGGAATATTTCTCACTGGAAGAAAGACTTAATCTTTTGAAGAGAGATGAGCG 1296
Qy |||||
989 AGTATCAGGTTTTCAGTTTATGCGCAGAAACCAAGATTAAGCTTTCACTTTGATGAGCAG 1048
Db |||||
1297 AGTATCAGAGATGGGAGTGTGTCAAGTCCACAGAGAAATCTTTTACCTTGGATGCTG 1356
Qy |||||
1049 ATTT 1052
Db |||||
1357 ACTT 1360

Search completed: September 26, 2005, 00:40:30
Job time : 470 secs


```
Db 301 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
QY 361 CGCTGTTTCTATGCTTCAATTTCTCATCAGAGATGTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGCTTCAATTTCTCATCAGAGATGTTCACTCAGAGATGTACAGTTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA 480
QY 481 ACCATGCCCTATGTTAAGAAAAGAGCAGATTTGGCCCTTGGCATGATAGCAGATAGAAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAAGAGCAGATTTGGCCCTTGGCATGATAGCAGATAGAAAA 540
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCCCTTTGCTGCTGTAGAGAGTGTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGGAAGAGTGGTGGCCCTTTGCTGCTGTAGAGAGTGTCTTCTCAGGA 600
QY 601 TCTTTTCTGCTATATTCTGGCTAAGAAAGAGAGGTCTTATGCCAGGACTCACATTTTTC 660
Db 601 TCTTTTCTGCTATATTCTGGCTAAGAAAGAGAGGTCTTATGCCAGGACTCACATTTTTC 660
QY 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA 720
Db 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA 720
QY 721 TACTTAGTAATAAGCCCTCAGAAGAAAGGTGAGGAGATCATTTGTGATGCTGTCAAA 780
Db 721 TACTTAGTAATAAGCCCTCAGAAGAAAGGTGAGGAGATCATTTGTGATGCTGTCAAA 780
QY 781 ATTGAGCAGAGTGTAAAGAGAGCTTGCAGTGTGCGCTCATTTGGAATGAATTCATT 840
Db 781 ATTGAGCAGAGTGTAAAGAGAGCTTGCAGTGTGCGCTCATTTGGAATGAATTCATT 840
QY 841 TTGATGAAACAGTACATTTGATGTTGTAGCTGACAGATTTACTTTGTGAACTTTGGAATTC 900
Db 841 TTGATGAAACAGTACATTTGATGTTGTAGCTGACAGATTTACTTTGTGAACTTTGGAATTC 900
QY 901 AAGGTTTTTTCAGGAGAAAATCCTTTTGTATTTATGAAAACATTTCTTTAGAGGAAAA 960
Db 901 AAGGTTTTTTCAGGAGAAAATCCTTTTGTATTTATGAAAACATTTCTTTAGAGGAAAA 960
QY 961 ACAAAATTTCTTTGAGAAACAGTTTTCAGAGTATCAGAGTGTGCGGTTTTCAGTTATGGCAGAAACC 1020
Db 961 ACAAAATTTCTTTGAGAAACAGTTTTCAGAGTATCAGAGTGTGCGGTTTTCAGTTATGGCAGAAACC 1020
QY 1021 ACAGATAACGTCCTTCACTTTGGATGCAGATTTT 1053
Db 1021 ACAGATAACGTCCTTCACTTTGGATGCAGATTTT 1053
```

RESULT 2

```
US-10-019-733-3
; Sequence 3, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 3
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-3

Query Match 100.0%; Score 1053; DB 4; Length 4955;
Best Local Similarity 100.0%; Pred. No. 2.1e-310;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGGCACC CGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 245 ATGGGGCACC CGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
QY 61 GACACCAACGAAAGTGAATAAAGTCAAAATCAAGAGCCACTCTCTAAGAAAGATTTCTCG 120
Db 305 GACACCAACGAAAGTGAATAAAGTCAAAATCAAGAGCCACTCTCTAAGAAAGATTTCTCG 364
QY 121 CGGTTTGTCTATCTTTTCCAATCCAGTACCTCATATTTTGGAAAATGTATAAACAGGCACAG 180
Db 365 CGGTTTGTCTATCTTTTCCAATCCAGTACCTCATATTTTGGAAAATGTATAAACAGGCACAG 424
QY 181 GCTTCTCTCTGGACAGCAGAGAGGTGCACCTTATCAAGAGATCTCCCTCACTGGAAACAG 240
Db 425 GCTTCTCTCTGGACAGCAGAGAGGTGCACCTTATCAAGAGATCTCCCTCACTGGAAACAG 484
QY 241 CTTAAAGCAGATGAGAGTACTTCTCATCTCTCACTCTTTAGCCTTTTTCAGCCAGTGTAT 300
Db 485 CTTAAAGCAGATGAGAGTACTTCTCATCTCTCACTCTTTAGCCTTTTTCAGCCAGTGTAT 544
QY 301 GGAATTTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAAGAGGTGCAGGTTCCAGAGGT 360
Db 545 GGAATTTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAAGAGGTGCAGGTTCCAGAGGT 604
QY 361 CGCTGTTTCTATGCTTTTCAAAATTTCTCATCAGAGATGTTCACTCAGAGATGTACAGTTTG 420
Db 605 CGCTGTTTCTATGCTTTTCAAAATTTCTCATCAGAGATGTTCACTCAGAGATGTACAGTTTG 664
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA 724
QY 481 ACCATGCCCTATGTTAAGAAAAGAGCAGATTTGGGCTTTGCGATGGATGAGATAGAAAA 540
Db 725 ACCATGCCCTATGTTAAGAAAAGAGCAGATTTGGGCTTTGCGATGGATGAGATAGAAAA 784
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCCCTTTGCTGCTGTAGAGAGTGTCTTCTCAGGA 600
Db 785 TCTACTTTTGGGGAAGAGTGGTGGCCCTTTGCTGCTGTAGAGAGTGTCTTCTCAGGA 844
QY 601 TCTTTTGTGCTATATTCTGGCTAAGAAAGAGAGGTCTTATGCCAGGACTCACTTTTTC 660
Db 845 TCTTTTGTGCTATATTCTGGCTAAGAAAGAGAGGTCTTATGCCAGGACTCACTTTTTC 904
QY 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTGTGATGTTTCCAA 720
Db 905 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTGTGATGTTTCCAA 964
QY 721 TACTTAGTAATAAGCCCTCAGAAGAAAGGTGCAGGAGATCATTTGTGATGCTGTCAAA 780
Db 965 TACTTAGTAATAAGCCCTCAGAAGAAAGGTGCAGGAGATCATTTGTGATGCTGTCAAA 1024
QY 781 ATTGAGCAGAGTGTAAAGAGAGCCCTTGGCCAGTTGGCCCTCATTTGGAATGAATTCATT 840
Db 1025 ATTGAGCAGAGTGTAAAGAGAGCCCTTGGCCAGTTGGCCCTCATTTGGAATGAATTCATT 1084
QY 841 TTGATGAAACAGTACATTTGATGTTGTAGCTGACAGATTTACTTTGTGAACTTTGGAATTC 900
Db 1085 TTGATGAAACAGTACATTTGATGTTGTAGCTGACAGATTTACTTTGTGAACTTTGGAATTC 1144
QY 901 AAGGTTTTTTCAGGAGAAAATCCTTTTGTATTTATGAAAACATTTCTTTTAGAGGAAAA 960
Db 1145 AAGGTTTTTTCAGGAGAAAATCCTTTTGTATTTATGAAAACATTTCTTTTAGAGGAAAA 1204
QY 961 ACAAAATTTCTTTGAGAAACAGTGTTCAGAGTATCAGCGTGTTCAGTGTATGCGAGAAC 1020
```

Db 1205 ACAAAATTTCTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTATGTCAGAGAAC 1264
Qy 1021 ACAGATAACGCTTCCACCTTGGATGACAGATTTT 1053
Db 1265 ACAGATAACGCTTCCACCTTGGATGACAGATTTT 1297

RESULT 3

US-10-019-733-12
; Sequence 12, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619W00P
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 12
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-12

Query Match 99.8%; Score 1051.4; DB 4; Length 1053;
Best Local Similarity 99.9%; Pred. No. 2.7e-310;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGGCGCACCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 ATGGCGCACCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Qy 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTCAAGAAAGAGTTCTGC 120
Db 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTCAAGAAAGAGTTCTGC 120
Qy 121 CGGTTTGTATCTTCCATCCAGTACCTCATATTTGGAATAATGTATAACAGGCACAG 180
Db 121 CGGTTTGTATCTTCCATCCAGTACCTCATATTTGGAATAATGTATAACAGGCACAG 180
Qy 181 GCTTCCTTCTGGACAGCAGAGAGTGCATTTATCAAGAGTCTCCCTCACTGGAACAG 240
Db 181 GCTTCCTTCTGGACAGCAGAGAGTGCATTTATCAAGAGTCTCCCTCACTGGAACAG 240
Qy 241 CTTAAGCAGATGAGAAGTACTTCTCATCTCAGATCTTAGCTTTTTCAGGCGAGTGAT 300
Db 241 CTTAAGCAGATGAGAAGTACTTCTCATCTCAGATCTTAGCTTTTTCAGGCGAGTGAT 300
Qy 301 GGAATGTGAATGAAATTTGGTGGAGCGCTTTTATGTCAGGAGGTGCAGGTTCCAGAGCT 360
Db 301 GGAATGTGAATGAAATTTGGTGGAGCGCTTTTATGTCAGGAGGTGCAGGTTCCAGAGCT 360
Qy 361 CGCTGTTTCTATGCTTTTCAATTTCTCATCAGAGTGTCTCATCAGAGATGACAGTTG 420
Db 361 CGCTGTTTCTATGCTTTTCAATTTCTCATCAGAGTGTCTCATCAGAGATGACAGTTG 420
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Qy 481 ACCATGCCCTTATGTTAAGAAAAAGCAGATTTGGGCTTTGGATGATGACAGATAGAAA 540
Db 481 ACCATGCCCTTATGTTAAGAAAAAGCAGATTTGGGCTTTGGATGATGACAGATAGAAA 540

Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGGCTGTGTAGAAAGAGTTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGGCTGTGTAGAAAGAGTTTTCTTCTCAGGA 600
Qy 601 TCTTTTGGCTCTATATTTCTGGCTAAAGAGAGAGGTCTTATGTCAGGATCTCATTTTTTC 660
Db 601 TCTTTTGGCTCTATATTTCTGGCTAAAGAGAGAGAGGTCTTATGTCAGGATCTCATTTTTTC 660
Qy 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCTACTGTGACTTTTGGCTTGGCTGTATGCCAA 720
Db 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCTACTGTGACTTTTGGCTTGGCTGTATGCCAA 720
Qy 721 TACTTAGTAAATAAGCCTTTCAGAGAAAGGGTCAGGAGATCATTTGTGATGCTGTCAA 780
Db 721 TACTTAGTAAATAAGCCTTTCAGAGAAAGGGTCAGGAGATCATTTGTGATGCTGTCAA 780
Qy 781 ATTGAGCAGGAGTTTTTAAACAGAACCTTGGCCCTCATTTGGAATGAATGCAAT 840
Db 781 ATTGAGCAGGAGTTTTTAAACAGAACCTTGGCCCTCATTTGGAATGAATGCAAT 840
Qy 841 TTGATGAACAGTACATTTGAGTTTCTAGCTGACAGATTTGTTGGAATTTCTCA 900
Db 841 TTGATGAACAGTACATTTGAGTTTCTAGCTGACAGATTTGTTGGAATTTCTCA 900
Qy 901 AAGTTTTTTCAGGCGAGAAATCCTTTTGAATTTTATGAAAAACATTTCTTTAGAGGAAA 960
Db 901 AAGTTTTTTCAGGCGAGAAATCCTTTTGAATTTTATGAAAAACATTTCTTTAGAGGAAA 960
Qy 961 ACAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTTGCAGTTATGCGAGAAACC 1020
Db 961 ACAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTTGCAGTTATGCGAGAAACC 1020
Qy 1021 ACAGATAACGCTTCCACCTTGGATGACAGATTTT 1053
Db 1021 ACAGATAACGCTTCCACCTTGGATGACAGATTTT 1053

RESULT 4

US-10-019-733-4
; Sequence 4, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619W00P
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 4
; LENGTH: 1081
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-4

Query Match 99.8%; Score 1051.4; DB 4; Length 1081;
Best Local Similarity 99.9%; Pred. No. 2.7e-310;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGGCGCACCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 20 ATGGCGCACCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 79
Qy 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAAGAAAGAGTTCTCGC 120
Db 80 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAAGAAAGAGTTCTCGC 139

```
Qy 121 CGGTTTGTCTCTTTTCAATCCAGTACCTCATATTTGGAAAATGTATAACAGGCACAG 180
Db |||||
Qy 140 CGGTTTGTCTCTTTTCAATCCAGTACCTCATATTTGGAAAATGTATAACAGGCACAG 199
Db |||||
Qy 181 GCTTCCCTTCTGGACAGCAGAGAGGTGCACTTATCAAGAGATCTCCCTCACTGAGAACAG 240
Db |||||
Qy 200 GCTTCCCTTCTGGACAGCAGAGAGGTGCACTTATCAAGAGATCTCCCTCACTGAGAACAG 259
Db |||||
Qy 241 CTTAAACAGATGAGAGTACTCTCATCTCTCACATCTTAGCCTTTTGGCAGCAGTGAT 300
Db |||||
Qy 260 CTTAAACAGATGAGAGTACTCTCATCTCTCACATCTTAGCCTTTTGGCAGCAGTGAT 319
Db |||||
Qy 301 GGAATTGTAATGAAAATTTGGTGAGCGCTTTAGTCAGAGAGTGCAGGTTCCAGAGGCT 360
Db |||||
Qy 320 GGAATTGTAATGAAAATTTGGTGAGCGCTTTAGTCAGAGAGTGCAGGTTCCAGAGGCT 379
Db |||||
Qy 361 CGCTGTTCTATGCTTTCAAATTTCTCATCGAGAAATGTTTCACTCAGAGATGACAGTTTG 420
Db |||||
Qy 380 CGCTGTTCTATGCTTTCAAATTTCTCATCGAGAAATGTTTCACTCAGAGATGACAGTTTG 439
Db |||||
Qy 421 CTGATAGACATTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATGAA 480
Db |||||
Qy 440 CTGATAGACATTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATGAA 499
Db |||||
Qy 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGGCCCTTGCATGATGATAGCAGATAGAAA 540
Db |||||
Qy 500 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGGCCCTTGCATGATGATAGCAGATAGAAA 559
Db |||||
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTCTGCTGTGAGAGAGTTTCTTCTCAGGA 600
Db |||||
Qy 560 TCTACTTTTGGGAAAGAGTGGTGGCTTTCTGCTGTGAGAGAGTTTCTTCTCAGGA 619
Db |||||
Qy 601 TCTTTTGTCTATATTTCTGCTTAAGAAAGAGGTCTTATGCCAGACTCCTTTTTC 660
Db |||||
Qy 620 TCTTTTGTCTATATTTCTGCTTAAGAAAGAGGTCTTATGCCAGACTCCTTTTTC 679
Db |||||
Qy 661 AATGAACTCATCAGCAGATGAGAGCTTCACTGTGACTTTTGTCTGCTCATGTTCCAA 720
Db |||||
Qy 680 AATGAACTCATCAGCAGATGAGAGCTTCACTGTGACTTTTGTCTGCTCATGTTCCAA 739
Db |||||
Qy 721 TACTTAGTAATAAGCCTTCCAGAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 780
Db |||||
Qy 740 TACTTAGTAATAAGCCTTCCAGAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 799
Db |||||
Qy 781 ATTGACAGAGGTTTAAACAGAGCCTTGCAGTGGCCTCATTTGGAATGATTCAT 840
Db |||||
Qy 800 ATTGACAGAGGTTTAAACAGAGCCTTGCAGTGGCCTCATTTGGAATGATTCAT 859
Db |||||
Qy 841 TTGATGAAAAGTACATTTGAGTTGTAGCTGACAGATTTACTTGTGGAATTTGGAATTTCA 900
Db |||||
Qy 860 TTGATGAAAAGTACATTTGAGTTGTAGCTGACAGATTTACTTGTGGAATTTGGAATTTCA 919
Db |||||
Qy 901 AAGGTTTTTTCAGGAGAAAATCCTTTTGATTTTATGAAAACATTTCTTTTGAAGGAAA 960
Db |||||
Qy 920 AAGGTTTTTTCAGGAGAAAATCCTTTTGATTTTATGAAAACATTTCTTTTGAAGGAAA 979
Db |||||
Qy 961 ACAAATTTTTCAGGAGAAAATCCTTTTGATTTTATGAAAACATTTCTTTTGAAGGAAA 1020
Db |||||
Qy 980 ACAAATTTTTCAGGAGAAAATCCTTTTGATTTTATGAAAACATTTCTTTTGAAGGAAA 1039
Db |||||
Qy 1021 ACAGATAACGCTTCACTTGGATGACAGATTTT 1053
Db |||||
Qy 1040 ACAGATAACGCTTCACTTGGATGACAGATTTT 1072
Db |||||
```

RESULT 5

US-09-962-665-9

; Sequence 9, Application US/09962665

; Patent No. 6537759

; GENERAL INFORMATION:

; APPLICANT: Stanton, Jf., Vincent P.

; TITLE OF INVENTION: POLYLPOLYGLUTAMATE SYNTHETASE GENE SEQUENCE

```
; TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING THE
; TITLE OF INVENTION: TREATMENT OF DISEASE
; FILE REFERENCE: 11926-015004
; CURRENT APPLICATION NUMBER: US/09/962,665
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 128,1464
; OTHER INFORMATION: n = g or a
; NAME/KEY: misc_feature
; LOCATION: 189
; OTHER INFORMATION: n = t or g
; NAME/KEY: misc_feature
; LOCATION: 524
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc_feature
; LOCATION: 1399
; OTHER INFORMATION: n = t or a
; NAME/KEY: misc_feature
; LOCATION: 1636,1738,2259
; OTHER INFORMATION: n = c or t
; OTHER INFORMATION: n = c or t
US-09-962-665-9
```

Query Match 55.0%; Score 579; DB 4; Length 2500;

Best Local Similarity 75.0%; Pred. No. 6.4e-166;

Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

```
Qy 89 ATGAGAGCCACTCTTAAGAAAGAGTTCTCGCGGTTTGTGTCATCTTTCCAAATCCAGTACC 148
Db |||||
Qy 397 AGATGAGCCCTGCTGAGAGAAACCCCGCGCTTTGTGTCATCTTTCCCATCCAGTACC 456
Db |||||
Qy 149 CTGATATTTGAAAATGTATAAACAGGCACAGGCTTCTCTTGGACAGCAGAGAGGTCG 208
Db |||||
Qy 457 ATGATATCTGCAGATGTATAAGAAAGGCAGAGGCTTCTTTTGGACCCCGAGGAGTTG 516
Db |||||
Qy 209 ACTTATCAAGGATCTCCCTCACTGGAAACAGCTTAAAGCAGATGAGAGTACTTCACT 268
Db |||||
Qy 517 ACCTCTCAAGAGCATTTGAGCACTCTGGAATCCCTGAAACCCCGAGGAGATATTTTATAT 576
Db |||||
Qy 269 CTCACTCTTAGCCTTTTTCAGCCACTGATGGAATGTAAATCGAAATTTGGTGGAGC 328
Db |||||
Qy 577 CCCATGTTCTGGCTTTCTTTGAGCAAGGATGGCATAGTAATGAAATCTTTGGTGGAGC 636
Db |||||
Qy 329 GCTTTAGTCAGAGGTCAGGTTTCAGAGGCTCGCTGTTTCTATGGCTTTTCAAAATTTCA 388
Db |||||
Qy 637 GATTTAGCAAGAGTTCAGATTACAGAAAGCCCGCTGTTTCTATGGCTTCCAAATTTGCCA 696
Db |||||
Qy 389 TCGAGATGTTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTATCATCAGAGATCCA 448
Db |||||
Qy 697 TGGAAAACATACATTTGAAATGTATAGTCTTCTTATTTGACACTTACATATAAAGATCCA 756
Db |||||
Qy 449 AGAAAGGGAATTTTATTTAATGCAATTTGAAACCATGCTTATGTAAGAAAAGCAG 508
Db |||||
Qy 757 AAGAAAGGGAATTTCTTCTCAATGCCATTTGAAACCATGCTTGTGTCTCAAGAAAGAGCAG 816
Db |||||
Qy 509 ATTGGGCTTTCGATGATAGATAGAAATCTACTTTTGGGAAAGAGTGGTGGCT 568
Db |||||
```

```
Db 817 ACTGGCCCTGGCTGGATGGGACAAAGGCTACCTATGCTGAACGTGTGTAGCCT 876
Qy 569 TTGCTGCTGTAGAGGATTTTCTCTCAGGATCTTTTGTGCTATATTTCTGGCTAAAGA 628
Db 877 TTGCTGCAGTGGAGGCAATTTCTTTCCGGTCTTTTGGCTCGATATTTGGCTCAAGA 936
Qy 629 AGAGAGGTCTTATGCCAGGACTCACTTTTCCAAATGAATCACTATCAGCAGAGATGAAGAC 688
Db 937 AACGAGGACTGATGCTGGCCTCACTTTTCTAAATGAATCTTATTAGCAGAGATGAGGCTT 996
Qy 689 TTCACTGTGACTTTGCTTGGCTGATGTTCCAAATCACTTAGTAATGAAGCTTTCAGNAGAA 748
Db 997 TACACTGTGATTTGCTTGGCTGATGTTCAAACTCGGTACACAAACCATCGGAGAGA 1056
Qy 749 GGGTCAGGAGATCAATTTGCTGATGCTGCAAAATTTGACGAGGATTTTAAACAGAGCCT 808
Db 1057 GAGTAAGAGAATAATATCAATGCTGTTCCGATAGACAGAGTTCCTCACTGAGGCT 1116
Qy 809 TGCAGTTGGCTCATTTGGAAATGAATTCATTTTGAATGAACAGACAGATTTGATG 868
Db 1117 TGCCTGTGAAGCTCAATTTGGGATGAATTCATCTAAATGAAGCAATACATTTGATG 1176
Qy 869 CTGACAGATTACTTGTGGACTTGGATCTCAAAAGTTTTTCAGGACGAAATCCTTTTG 928
Db 1177 CAGACAGACTTATGCTGGAAGTGTGTTTACCAAGGTTTTCAGAGTAGAGAACCCATTTG 1236
Qy 929 ATTTTATGGAACAAATTTCTTGAAGGAAACAAATTTCTTTGAGAAACGAGTTTCAG 988
Db 1357 ACTT 1360
```

RESULT 6

```
US-09-963-333-9
; Sequence 9, Application US/09963333
; Patent No. 6664062
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: THYMIDINE SYNTHASE GENE SEQUENCE VARIANCES
; TITLE OF INVENTION: HAVING UTILITY IN DETERMINING THE TREATMENT
; TITLE OF INVENTION: OF DISEASE
; FILE REFERENCE: 11926-015002
; CURRENT APPLICATION NUMBER: US/09/963,333
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 128,1464
; OTHER INFORMATION: n = g or a
; NAME/KEY: misc_feature
; LOCATION: 189
```

```
; OTHER INFORMATION: n = t or g
; NAME/KEY: misc_feature
; LOCATION: 524
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc_feature
; LOCATION: 1399
; OTHER INFORMATION: n = t or a
; NAME/KEY: misc_feature
; LOCATION: 1636, 1738, 2259
; OTHER INFORMATION: n = c or t
US-09-963-333-9
```

```
Query Match 55.0%; Score 579; DB 4; Length 2500;
Best Local Similarity 75.0%; Pred. No. 6.4e-166;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;
```

```
Qy 89 ATGAAGACCCTCTCTAAGAAAGATTCTCGCCGGTTTGTCTATCTTTCCAAATCCAGTACC 148
Db 397 AGGATGAGCCGCTGCTGAGAGAAAACCCCGCCGCTTTGTCTATCTTTCCCATCGAGTACC 456
Qy 149 CTGATATTTTGAAATATGTAACAGGACACAGGCTTCTCTCTGGACAGCAGAGAGGCTCG 208
Db 457 ATGATATCTGGCAGATGATAGAAAGGAGAGGCTTCTTTTGGACCCCGGAGGAGTTG 516
Qy 209 ACTTATCAAGGATCTCCCTCACTGGAACAAAGCTTAAAGCAGATGAGAAATGATCTCATCT 268
Db 517 ACCTCTNAAGGACATTCAGCACTGGGATCCCTGAAACCCGAGGAGATATTTTATAT 576
Qy 269 CTCACATCTTAGCCCTTTTTCAGCCAGTGAATGTAATGAAATTTGGTGGAGC 328
Db 577 CCCATGTTCTGGCTTTCTTTGCGCAGCAAGCATGGCATAGTAATGAAACTTTGGTGGAGC 636
Qy 329 GCTTTAGTCAGAGGTGCGAGTTCAGAGGCTCGCTGTTCTATGGCTTTCAAAATCTCA 388
Db 637 GATTTAGCAAGAGTTCAGATTAAGAAAGCCCGCTGTTCTTATGGCTTTCCAAATTTGCCA 696
Qy 389 TCGAATATGTTCACTCAGAGATGTACAGTTGCTCATAGACACTTACATCAGAGATCCCA 448
Db 697 TGGAAACATACATCTGAAATGATATGTTCTTTATGACACTTACATPAAAGATCCCA 756
Qy 449 AGAAAGGGAAATTTTATTTAATGCAATTTGAAACCATGCCCTATGTTAAGAAAAAGCAG 508
Db 757 AAGAAAGGGAAATTTCTCTTCAATGCCATTTGAAACCATGCCCTTTGTGTCAAGAAAGAGCAG 816
Qy 509 ATTGGCCCTTCGATGATGATAGATGAAATCTACTTTTGGGGAAGAGAGTGGTGCCT 568
Db 817 ACTGGCCCTTGGCTGGATTTGGGACAAAGAGGCTACCTATGTTGTAACGTTGTAGCCT 876
Qy 569 TTGCTGCTGTAGAAGGAGTTTCTTCTCAGGATCTTTTGTGCTGCTATATTTCTGCTAAAGA 628
Db 877 TTGCTGAGTGAAGGAGGCAATTTCTTTCCGGTCTTTTTCGGTCTGATATTTCTGGCTCAAGA 936
Qy 629 AGAGAGGTCTTATGCCAGGACTCACTTTTTCATGAACTCATCAGCAGAGATGAAGGAC 688
Db 937 AACGAGGACTGATGCTGGCTCACTTTTCTAATGAATTTATAGCAGAGATGAGGCTT 996
Qy 689 TTCACTGTGACTTTGCTTGGCTGATGTTCCAAATCACTTAGTAATGAAGCTTTCAGAGAAA 748
Db 997 TACACTGTGATTTTGGCTTGGCTGATGTTCAAACTCGGTACACAAACCATCGGAGAGA 1056
Qy 749 GGGTCAGGAGATCAATTTGCTGATGCTGCAAAATTTGACGAGGATTTTAAACAGAGCCT 808
Db 1057 GAGTAAGAGAATAATATCAATGCTGTTCCGATAGAACAGAGGTTCTCTCACTGAGGCT 1116
Qy 809 TGCCAGTTGGCTCATTTGGAAATGAATTCATTTTGAATGAACAGTACATTTGATTTGTAG 868
Db 1117 TGCCTGTGAAGCTCAATTTGGGATGAATTCATCTAAATGAAGCAATACATTTGATTTG 1176
Qy 869 CTGACAGATTACTTGTGGAACTTGGATTTCTCAAAAGTTTTTCAGGACGAAATCTTTTG 928
Db 1177 CAGACAGACTTATGCTGGAACTGGGTTTATAGCAAGGTTTTCAGAGTAGAGAACCCATTTG 1236
Qy 929 ATTTTATGGAACAAATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACGAGTTTCAG 988
```

Db 1237 ACTTATGGAGATATTTCTACGTGAAGAAAGACTAACTTTTGGAGAGAGACTAGGCG 1296
Qy 989 AGTATCAGCGTTTGGCAGTTATGCGCAGAAACACAGATAAGCTTCCACCTTGGATGACG 1048
Db 1297 AGTATCAGAGATGGGAGTGTGATGTCAGATCCACAGAGAATTTCTTTACCTTGGATGCTG 1356
Qy 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 7

US-09-962-677-9

; Sequence 9, Application US/09962677

; Patent No. 6759200

; GENERAL INFORMATION:

; APPLICANT: Stanton, Jr., Vincent P.

; TITLE OF INVENTION: THYMIDINE PHOSPHORYLASE GENE SEQUENCE

; TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING

; TITLE OF INVENTION: THE TREATMENT OF DISEASE

; FILE REFERENCE: 11926-015003

; CURRENT APPLICATION NUMBER: US/09/962,677

; CURRENT FILING DATE: 2001-09-24

; PRIOR APPLICATION NUMBER: 09/658,659

; PRIOR FILING DATE: 2000-09-08

; PRIOR APPLICATION NUMBER: 09/596,033

; PRIOR FILING DATE: 2000-06-15

; PRIOR APPLICATION NUMBER: 09/357,743

; PRIOR FILING DATE: 1999-07-20

; PRIOR APPLICATION NUMBER: 09/357,024

; PRIOR FILING DATE: 1999-07-19

; PRIOR APPLICATION NUMBER: 60/093,484

; PRIOR FILING DATE: 1998-07-20

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 9

; LENGTH: 2500

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: 128..1464

; OTHER INFORMATION: n = g or a

; NAME/KEY: misc_feature

; LOCATION: 189

; OTHER INFORMATION: n = t or g

; NAME/KEY: misc_feature

; LOCATION: 524

; OTHER INFORMATION: n = c or g

; NAME/KEY: misc_feature

; LOCATION: 1399

; OTHER INFORMATION: n = t or a

; NAME/KEY: misc_feature

; LOCATION: 1636..1738, 2259

; OTHER INFORMATION: n = c or t

; US-09-962-677-9

Query Match

Best Local Similarity 55.0%; Score 579; DB 4; Length 2500;

Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

Qy 89 ATGAGAGCCACTCCTAAGAAAGAGTCTGCGCGGTTGTGTCATCTTTCCATCCAGTACC 148
Db 397 AGGATGAGCGCTGCTGAGAGAAAACCCCGCGCTTTGTGTCATCTTCCCATCCAGTACC 456
Qy 149 CTGATATTTGAAAATGTATAAACAGGCACAGGCTTCTTCTTGACAGCAGAGGTGCG 208
Db 457 ATGATATCTGGCAGATGATAGAGGAGAGGCTTCTTTTGGACCGCGAGGAGGTG 516
Qy 209 ACTTATCAAGAGTCTCCCTCCTCCTGGAACAGCTTAAAGCAGAGATGAGAAGTACTTCATCT 268
Db 517 ACCTCTCNAAGGACATTCAGCAGCTGGGAATCCCTGAAACCGGAGAGAGATATTTTATAT 576

Qy 269 CTACATCTTAGCCCTTTTTCAGCCAGTGATGAATTTGTAATGAAAATTTTGGTGAGC 328
Db 577 CCATGTTCTGGCTTTTTCAGCAAGCGATGGCAATAGTAATGAAAATTTTGGTGAGC 636
Qy 329 GCTTTAGTCAGAGAGTTCAGAGTTCAGAGGCTCGCTGTTTCTATGGCTTTCAAATTTCTCA 388
Db 637 GATTTAGCCAAAGAGTTTCAGATTACAGAAAGCCGCTGTTTCTATGGCTTCCAAATTTGCCA 696
Qy 389 TCAGAAATGTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAGATCCCA 448
Db 697 TGGAAAACATACATTTCTGAAATGTATAGTCTTCTTATTGACACTTACATAAAGATCCCA 756
Qy 449 AGAAAAGGGAATTTTATTAAATGCAATTTGAAACCATGCCCTATGTTAAGAAAAGAGCAG 508
Db 757 AAGAAAGGGAATTTCTCTTAATGCCATGAAACGATGCCCTGTGTCAAGAAAGAGGAG 816
Qy 509 ATTGGGCTTTCGATAGATAGATAGAAATCTACTTTTGGGGAAGAGTGGTGGCT 568
Db 817 ACTGGGCTTTCGCTGGATTCGGGACAAAGAGGCTACCTATGTTGAACGTTGTAGCCT 876
Qy 569 TTGCTGCTGTAAGAGAGTTTTCTTCTCAGGATCTTTTGTCTGCTATATTTCTGGCTAAAGA 628
Db 877 TTGCTGCAAGTGAAGGCAATTTTCTTTCCGGTCTTTTTCGGTCTGATATTTCTGGCTCAAG 936
Qy 629 AGAGAGGTCTTATGCCAGGACTCACTTTTCCAAATGAACTCATCAGCAGAGATGAAGGAC 688
Db 937 AACGAGGACTGATGCCCTGCCCTCAATTTCTAATGAACTTATTAGCAGAGATGAGGTT 996
Qy 689 TTCACTGTGACTTTTGTGCTGATGTTCCAAATCTTAGTAATTAAGCCTTCAGAGAGAA 748
Db 997 TACACTGTGATTTTGTGCTGATGTTTCAAAACACCTGTTACACAAACCATCGGAGGAGA 1056
Qy 749 GGGTCAGGAGATCAITTTGATGCTGTGCTGCAAAATTCAGCAGAGATTTTAAACAGAGCCT 808
Db 1057 GAGTAAGAGAAATTAATTAATCAATGCTGTTCGGATAGAAACAGGAGTTCTCTCACTGAGCCT 1116
Qy 809 TGCCAGTTGGCTCATTTGGAATGAATTCATTTTGCATGAAACAGTACATGAGTTTGTAG 868
Db 1117 TGCCTGTGAAGCTCATTGGGATGAATTCATCTTAATGAGCAATACATTTGATTTGTGG 1176
Qy 869 CTGACAGATTACTTTGTGGAATCTGGAATTTCTCAAAGTTTTCAGGCAGAAAATCCTTTTG 928
Db 1177 CAGACAGACTTATGCTGGAATCTGGGTTTTCAGAGGTTTTCAGAGTAGAGAACCATTTG 1236
Qy 929 ATTTTATGGAACATTTCTTTTAGAGGAAAACAAATTTCTTTTGAAGAACAGATTTTCA 988
Db 1237 ACTTTATGGAGATATTTTCACTGGAAGGAAAGACTTAACCTTCTTTGAGAAAGAGTAGGCG 1296
Qy 989 AGTATCAGCGTTTTCAGTTATGCGAGAAACCAAGATTAACGCTTTCACCTTGGATGCG 1048
Db 1297 AGTATCAGAGATGGGAGTGTGTCAGTCCAAACAGAGAAATTTCTTTTACCTTTGGATGCTG 1356
Qy 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 8

US-09-949-016-2025

; Sequence 2025, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: C1001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

;; PRIOR APPLICATION NUMBER: 60/231,498
;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 2025
;; LENGTH: 2479
;; TYPE: DNA
;; ORGANISM: Human
US-09-949-016-2025

Query Match 54.9%; Score 578.4; DB 4; Length 2479;
Best Local Similarity 75.0%; Pred. No. 9.6e-166;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

```
Qy 89 ATGAAGAGCCACTCTCTAAGAAAGAGTCTCGCCGGTCTTGTATCTTTCCAAATCCAGTACC 148
Db |||||
Qy 397 AGGATGAGCCGCTGCTGAGAGAAACCCCGCGCTTGTATCTTTCCCATCCAGTACC 456
Db |||||
Qy 149 CTGATATTGGAATAATGATATAACAGGCAAGGCTTCTTCTGGACAGCAGAGAGTGG 208
Db |||||
Qy 457 ATGATATCTGGCAGATGATATAAGAGGCGAGAGGCTTCTTTGGACCGCGAGAGTGG 516
Db |||||
Qy 209 ACTTATCAAGAGTCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAGTACTTCTATCT 268
Db |||||
Qy 517 ACCTCTCAAGGACATTCAGCACTGGGAATCCCTGAACCCGAGGAGAGATATTTATAT 576
Db |||||
Qy 269 CTCACATCTTAGCCCTTTTTCAGCCAGTGATGAATTTAAATCAAAATTTGGTGAGC 328
Db |||||
Qy 577 CCCATGTTCTGGCTTTCTTTCAGCAAGCGATGGCATAGTAATGAANAATTTGGTGAGC 636
Db |||||
Qy 329 GCTTTAGTTCAGGAGGTGCGAGTTCAGAGGCTCGCTGTTTCTATGGCTTTCAAAATTTCA 388
Db |||||
Qy 637 GATTTAGCAAGAGTTCAGATTACAGAGCCGCTGTTTCTATGGCTTCCAAATTTGCA 696
Db |||||
Qy 389 TCGAAGATGTTCACTCAGAGATGACAGTGTGCTGATAGACACTTACATCAGAGATCCCA 448
Db |||||
Qy 697 TGGAAACATACATCTGAAATGTATAGTCTTCTATTGACACTTACATAAAGATCCA 756
Db |||||
Qy 449 AGAAAGGGATTTTATTAATCAATGGAACCAATGCAATGCTTATGTTAAGAAAGAGCAG 508
Db |||||
Qy 757 AAGAAAGGGAATTTCTCTTCAATGCCATTTGAACGATGCTTGTCTCAAGAGAGGCGAG 816
Db |||||
Qy 509 ATTGGGCTTCGATGATAGACATAGAAATCTACTTTTGGGAAAGAGTGGTGCT 568
Db |||||
Qy 817 ACTGGGCTTCGCTGATGAGGGAAGAGGCTACTATGTTGGAAGCTGTTGAGCT 876
Db |||||
Qy 569 TTGCTGCTGTAGAGGAGTCTTCTCTCAGGATCTTTGCTGCTATATTTCTGGCTAAGA 628
Db |||||
Qy 877 TTGCTGAGTGGAGGCAATTTCTTTCCGGTCTTTTGGCTGATATTTCTGGCTCAAGA 936
Db |||||
Qy 629 AGAGAGTCTTATGCCAGGACTCACTTTTCCAAATGAATCATCAGCAGAGATGAAGAC 688
Db |||||
Qy 937 AACGAGGACTGATGCTGCGCTCAATTTCTAATGAATTTATAGCAGAGATGAGGTT 996
Db |||||
Qy 689 TTCACTGTGATTTGCTGCTGATGTTTCCAAATCTAGTAATTAAGCTTCAAGAGAA 748
Db |||||
Qy 997 TACACTGTGATTTTGGCTTGGCTGAATTTCAACACCTGGTACACAAACCATCGGAGGA 1056
Db |||||
Qy 749 GGGTCAGGAGATCATTTGTTGATCTCTCAAAATTTGAGCAGGAGTTTTAAACAGAGCT 808
Db |||||
Qy 1057 GAGTAAGAGAAATTAATTAATCAATGCTGTTCCGATAGAACAGGAGTTCCTCACTAGGCT 1116
Db |||||
Qy 809 TGCAGTGTGGCTCATTTGGAAATGAATTTGCAATTTTGAATGAACAGTACATTCAGTTTGG 868
Db |||||
Qy 1117 TGCCTGTGAAGCTCATTTGGGATGAATTTGCACTCTAATGAAGCAATACATTCAGTTTGG 1176
Db |||||
Qy 869 CTGACATTTACTTGTGGAATTTGATCTTCAAGGTTTTCAGCAGCAAAATTCCTTTG 928
Db |||||
Qy 1177 CAGACAGACTTATGCTGGAATGCTGGTTTTCAGCAAGGTTTTCAGAGTAGAGAACCATTTG 1236
Db |||||
Qy 929 ATTTTATGGAACATTTCTTTAAGAGAAACAAATTTCTTTGAGAAACGAGTTTCAG 988
Db |||||
Qy 1237 ACTTTATGGAATATTTCTCTGAGAGAAAGACTAATCTTTTGGAGAGAGATAGCGC 1296
Db |||||
```

```
Qy 989 AGTATCAGCGTTTTCAGTATTGCGAGAAACACAGATAACGTTCTTCCCTTTGATGCG 1048
Db |||||
Qy 1297 AGTATCAGAGGATGGAGTGAATGTCAGTCAAGTCCAACAGAGATTCTTTTACCTTGGATGCTG 1356
Db |||||
Qy 1049 ATTT 1052
Db |||||
Qy 1357 ACTT 1360
Db |||||
```

RESULT 9

US-09-023-655-1370
; Sequence 1370, Application US/09023655
; Patent No. 6607879

;; GENERAL INFORMATION:
;; APPLICANT: Cocks, Benjamin G.
;; APPLICANT: Susan G. Stuart
;; APPLICANT: Jeffrey J. Seilhamer
;; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
;; NUMBER OF SEQUENCES: 1508
;; CORRESPONDENCE ADDRESS:
;; ADDRESSER: INCYTE PHARMACEUTICALS, INC.
;; STREET: 3174 PORTER DRIVE
;; CITY: PALO ALTO
;; STATE: CALIFORNIA
;; COUNTRY: USA
;; ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/023,655
FILING DATE: HERewith
CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:

NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0001 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1370:

SEQUENCE CHARACTERISTICS:

LENGTH: 2500 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g36154
US-09-023-655-1370

Query Match 54.9%; Score 578.4; DB 4; Length 2500;
Best Local Similarity 75.0%; Pred. No. 9.7e-166;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

```
Qy 89 ATGAAGAGCCACTCTCTAAGAAAGAGTCTCGCCGGTCTTGTATCTTTCCAAATCCAGTACC 148
Db |||||
Qy 397 AGGATGAGCCGCTGCTGAGAGAAACCCCGCGCTTGTATCTTTCCCATCCAGTACC 456
Db |||||
Qy 149 CTGATATTGGAATAATGATATAACAGGCAAGGCTTCTTCTGGACAGCAGAGAGTGG 208
Db |||||
Qy 457 ATGATATCTGGCAGATGATATAAGAGGCGAGAGGCTTCTTTTGGACCGCGAGAGGTTG 516
Db |||||
Qy 209 ACTTATCAAGAGTCTCCCTCACTGGAACAGCTTAAAGCAGATGAGAGTACTTCTATCT 268
Db |||||
```

Db 517 ACCTCTCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT 576
Qy 269 CTCACTCTTAGCTTTTTCAGCCAGTGTGAATTTGTAATAAGAAATTTGCTGGAGC 328
Db 577 CCNATGTTCTGGCTTTCTTTCAGCAAGCGATGCAATAGTAATAAGAAATTTGCTGGAGC 636
Qy 329 GCTTTAGTCAGGAGGTTCAGAGTTCAGAGGCTCGCTGTCTTATGGCTTTCAAAATTTCTCA 388
Db 637 GAATTTAGCAAGAGTTTCAGATTACAGAGCCCGCTGTCTTATGGCTTCCAAATTTGCCA 696
Qy 389 TCGAAGATGTTCACTCAGAGATGTACAGTTGTGCTGTAGATAGACACTATCAGAGATCCCA 448
Db 697 TGGAAACATACATTTCTGAAATGTATGTCTTCTTATTTGACACTTACATTAAGATCCCA 756
Qy 449 AGAAAGGGGAATTTTATTAATCAATTTGAATCAATGCAATGCGCTATGTTAAGAAAGGAGC 508
Db 757 AAGAAAGGGAATTTCTCTTCAATGCCATTTGAACGATGCTTGTCTAAGAAAGGAGC 816
Qy 509 ATTGGGCTTTGCGATGATAGCAGATAGAAATCTACTTTTGGGGAAGAGTGTGGCCT 568
Db 817 ACTGGGCTTTGCGCTGATTTGGGACAAAGAGGCTACCTATGTTGAACGTGTTGTAGCCT 876
Qy 569 TTGCTGCTGTAGAGAGTTTCTTCTCAGAGATCTTTTGTGCTGTATATTTCTGGCTAAAGA 628
Db 877 TTGCTGAGTGGAGGCAATTTCTTTTCCGGTCTCTTTTGGCTCGATATTTCTGGCTCAAGA 936
Qy 629 AGAGAGTCTTATGCCAGGACTCACTTTTCCAAATGAACTCATCAGCAGAGATGAAGGAC 688
Db 937 AACGAGACTGATGCCCTGCGCTCAATTTTCTAATGAACCTTATTAGCAGAGATGAGGGTT 996
Qy 689 TTCACTGTGACTTTGCTGCTGATGTTTCCAAATCTTAGTAATAAGCCTTTCAAGAGAA 748
Db 997 TACACTGTGATTTTGTCTGCTGATGTTTCAACACCTGTTACACAACCATCGAGGAGA 1056
Qy 749 GGGTCAGGAGATCAATTTGTGATGCTGTCAAAATGAGCAGAGTTTTAAAGAGCCT 808
Db 1057 GAGTAAGAGAAATTAATTAATCAATGCTGTTCCGATAGAACAGGAGTTCTCACTGAGGCT 1116
Qy 809 TGCAGTGGCCTCAATTTGGAATGAAATTTGATGAAACAGTACATTTGATGTTGTAG 928
Db 1177 TGCCTGTGAAGCTCAATTTGGGATGAATTCGACTCTTAATGAAGCAATACATTTGATTTGG 1176
Qy 929 ATTTTATGGAACATTTCTTTTAGAGGAAACAAATTTCTTTTGAAGAACGAGTTTTCAG 988
Db 1237 ACTTTATGGAGATATTTTCACTGGAAGGAAAGACTAACTTCTTTGAGAAGAGATGAGCG 1296
Qy 989 AGTATCAGCGTTTTCAGATTATGGCAGAAACCAAGATACGTTCTCACTTGGATGCG 1048
Db 1297 AGTATCAGAGATGGGAGTGTGTAAGTCCAAAGTCCAAAGAGAAATTTCTTTACCTGGATGCTG 1356
Qy 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 10
US-09-949-016-145
; Sequence 145, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 145
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-145

Query Match 54.9%; Score 578.4; DB 4; Length 2500;
Best Local Similarity 75.0%; Pred. No. 9.7e-166;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

Qy 89 ATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGCGGTTTGTCTATCTTCCAAATCCAGTACC 148
Db 397 AGGATGAGCCGCTCTGTAGAGAAACCCCGCCGCTTTGTCTATCTTCCCCTACGAGTACC 456
Qy 149 CTGATATTTGGAAATGTATAACACAGGCACAGGCTTCTCTTGGACAGCAGAGAGAGGTGG 208
Db 457 ATGATATCTGGCAGATGTATAAGAGGCGAGAGGCTTCTTTTGGACCCGCGAGAGGTGG 516
Qy 209 ACTTATCAAGAGATCTCCCTCACTGGAACAAAGCTTTAAAGCAGATGAGAGTACTTCACT 268
Db 517 ACCTCTCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT 576
Qy 269 CTCACTCTTAGCCCTTTTTCAGCCAGTGTAGTAATTTGTAATCAAAATTTGGTGGAGC 328
Db 577 CCCATGTTCTGGCTTTCTTTTCGCAAGCGATGGCATAGTAATAATGAAATCTTTGGTGGAGC 636
Qy 329 GCTTTAGTCAGAGGTGCGAGTTCAGAGGCTGCGTGTCTTCTATCGCTTTCAAAATTTCTCA 388
Db 637 GATTTAGCCAGAGAGTTCAGATTACAGAGCCCGCTGTTTCTATGGCTTCCAAATTTGCCA 696
Qy 389 TCGAAGATGTTCACTCAGAGATGTACAGTTTGTGCTGATAGACACTTACATCAGAGATCCCA 448
Db 697 TGGAAACATACATTCGAAATGTATAGTCTTCTTATTTGACACTTACATAAAGATCCCA 756
Qy 449 AGAAAGGGAATTTTATTAATGCAATGTAACCATGCCCTATGTTAAGAAAGAGCAG 508
Db 757 AAGAAAGGGAATTTCTCTTCAATGCCATTTGAAACCGATGCTTGTCTCAAGAAAGAGGAGC 816
Qy 509 ATTGGGCTTTCGATGATAGCAGATGAAATCTACTTTTGGGGAAGAGTGTGGCCT 568
Db 817 ACTGGCCTTTCGCTGGAATGGGACAAAGAGGCTTACCTATGTTGAACTGTTGTAGCCT 876
Qy 569 TTGCTGCTGTAGAGGAGTTTCTTCTCAGGATCTTTTGTGCTGTATATTTCTGGCTAAAGA 628
Db 877 TTGCTGCAAGTGAAGGCAATTTCTTTCCGGTCTTTTTCGGTCTGATATTTCTGGCTCAAGA 936
Qy 629 AGAGAGTCTTATGCGCAGGACTCACTTTTTCGAATGAACTCATCAGCAGAGATGAAGGAC 688
Db 937 AACGAGACTGATGCCCTGCGCTCAATTTTCTAATGAACCTTATTAGCAGAGATGAGGGTT 996
Qy 689 TTCACTGTGACTTTGCTGCTGATGTTTCCAAATCTTAGTAATAAGCCTTTCAAGAGAA 748
Db 997 TACACTGTGATTTTGTCTGCTGATGTTTCAACACCTGTTACACAACCATCGAGGAGA 1056
Qy 749 GGGTCAGGAGATCAATTTGTGATGCTGTCAAAATGAGCAGAGTTTTAAAGAGCCT 808
Db 1057 GAGTAAGAGAAATTAATTAATCAATGCTGTTCCGATAGAACAGGAGTTCTCACTGAGGCT 1116
Qy 809 TGCAGTGGCCTCAATTTGGAATGAAATTTGATGAAACAGTACATTTGATGTTGTAG 868
Db 1177 TGCCTGTGAAGCTCAATTTGGGATGAATTCGACTCTTAATGAAGCAATACATTTGATTTGG 1176
Qy 869 CTGACAGATTTACTTGTGGAACCTTGGATTTCTCAAAGTGTTCAGGAGGATTAATCTTTTG 928
Db 1177 CAGACAGACTTATGCTGGAACCTGAGTTTTCAGAGGTTTTCAGAGTAGAACCATTTG 988
Qy 929 ATTTTATGGAACATTTCTTTTAGAGGAAACAAATTTCTTTTGAAGAACGAGTTTTCAG 988
Db 1237 ACTTTATGGAGATATTTTCACTGGAAGGAAAGACTAACTTCTTTGAGAAGAGATGAGCG 1296
Qy 989 AGTATCAGCGTTTTCAGATTATGGCAGAAACCAAGATACGTTCTCACTTGGATGCG 1048
Db 1297 AGTATCAGAGATGGGAGTGTGTAAGTCCAAAGTCCAAAGAGAAATTTCTTTACCTGGATGCTG 1356
Qy 1049 ATTT 1052
Db 1357 ACTT 1360

Db 1237 ACTTATCGAGATATTTCTACCTGGAGGAAGACTAACTCTTTTGGAGAGAGTAGGCG 1296
Qy 989 AGTATCAGCGTTTTCAGTTATGCGAGAAACACAGATAAGCTCTTCACTTGGATGCAG 1048
Db 1297 AGTATCAGAGTAGGAGTAGTCAAGTCCAACAGAGAAATCTTTTACCTTGGATGCTG 1356
Qy 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 11

US-08-905-223-125
; Sequence 125, Application US/08905223
; Patent No. 622029
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste D.
; APPLICANT: Duclert, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: 5' ESTs FOR SECRETED PROTEINS
; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Win95
SOFTWARE: Word

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/905,223
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Israel, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 125:

SEQUENCE CHARACTERISTICS:
LENGTH: 481 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: DOUBLE
TOPOLOGY: LINEAR
MOLECULE TYPE: CDNA
ORIGINAL SOURCE:
ORGANISM: Homo Sapiens
TISSUE TYPE: Brain
FEATURE:
NAME/KEY: sig peptide
LOCATION: 41...343
IDENTIFICATION METHOD: Von Heijne matrix
OTHER INFORMATION: score 4.4
OTHER INFORMATION: seq ISHILAFFAASDG/IV

US-08-905-223-125

Query Match 41.9%; Score 441; DB 3; Length 481;
Best local Similarity 100.0%; Pred. No. 3.2e-124;
Matches 441; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGGCGACCGGAAAGCGGCGGCTGGATCAGATGAGATCATCTTCA 60
Db 41 ATGGGCGACCGGAAAGCGGCGGCTGGATCAGATGAGATCATCTTCA 100
Qy 61 GACACCAAGAAAGTGAATAAAGTCAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 101 GACACCAAGAAAGTGAATAAAGTCAAGAGCCACTCTTAAGAAAGAGTTCTCGC 160

Qy 121 CGTTTGTCTATCTTTCCATCCAGTACCTGATATTTGAAAATGTATAACAGGCACAG 180
Db 161 CGTTTGTCTATCTTTCCATCCAGTACCTGATATTTGAAAATGTATAACAGGCACAG 220
Qy 181 GCTTCTTCTGACAGCAGAGAGGTGCACTTATCAAGAGATCTCCCTCACTGGAAACAAG 240
Db 221 GCTTCTTCTGACAGCAGAGAGGTGCACTTATCAAGAGATCTCCCTCACTGGAAACAAG 280
Qy 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCATCTTAGCCCTTTTTCAGCCAGCTGAT 300
Db 281 CTTAAAGCAGATGAGAAGTACTTCTCATCTCATCTTAGCCCTTTTTCAGCCAGCTGAT 340
Qy 301 GGAATTGTAATGAAAATTTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGTTCAGAGGCT 360
Db 341 GGAATTGTAATGAAAATTTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGTTCAGAGGCT 400
Qy 361 CGCTGTTCTATGCGCTTTCAAATTTCTCATCGAAGATGTTCACTCAGAGATGATACAGTTTG 420
Db 401 CGCTGTTCTATGCGCTTTCAAATTTCTCATCGAAGATGTTCACTCAGAGATGATACAGTTTG 460
Qy 421 CTGATAGACACTTACATCAGA 441
Db 461 CTGATAGACACTTACATCAGA 481

RESULT 12

US-08-307-499-1/c
; Sequence 1, Application US/08307499
; Patent No. 5651972
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: U.S.A.
; ZIP: 32606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/307,499
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/908,241
FILING DATE: 1-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/908,630
FILING DATE: 29-JUN-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/342,212
FILING DATE: 21-APR-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UF35-1.FWCCI
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:


```

; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 138..1460
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2456..2659
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3070..3330
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3356..4180
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 4392..5894
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 6171..6398
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 6447..6875
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 6928..7431
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 7454..7858
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 7895..8155
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 8215..8682
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 8715..9539
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 9562..10272
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 10316..11908
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 11971..12780
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 12829..13107
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 13149..14171
; US-08-307-499-14

Query Match 40.8%; Score 429.8; DB 1; Length 14176;
Best Local Similarity 66.2%; Pred. No. 5.9e-120;
Matches 651; Conservative 0; Mismatches 327; Indels 5; Gaps 2;

QY 71 AAGTGAATAAAGTCAAAATGAAGAGCCACTCTAAGAAAGAGTCTCGCGGTTTGTC 130
Db 3336 AAAATGATTTTATACAGATGAGGCTATCTTCAAGACTCTGATCTAGGTTGTTA 3395

QY 131 TCTTTCCCAATCCAGTACCTGATATTTGGAATAATGTATATAACAGGACAGGCTTCCTTCT 190
Db 3396 TTTTCCCTATTAAAGTATCATGATATCTGGAATAATGTATATAACATCATGTCGGAAGTTT 3455

QY 191 GGACGACGAGAAGGTCGACTATATCAAGGATCTCCCTCACTGGNAACAGCTTAAAGCAG 250

3456 GGACCGTTGAAGAAGTAGATTTATCAAAAGATTTAGATGATTGGGATAAATTAACATAAAG 3515
251 ATGAGAAGTACTTCTCATCTCTCATCTTCTAGCCCTTTTTCAGCCAGCTGATGAATTTGTAA 310
3516 ACGAAAAATACTTTTATAAAACATATACCTAGCAATTTTTCATCTAGTATGTTGTAA 3575
311 ATGAAAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCTCGCTGTTTCT 370
3576 ATGAGAAATTTAGCGGAAGATTTTATGTGATGACAGTGTTCAGAGGACGATGTTTCT 3635
371 ATGGCTTTCAAATCTCATGAGAAATGTTCACTCAGAGATGACAGTGTTCGATGATGACA 430
3636 ATGGATTTCAAATAGCTATGGAATAATTCATTCAGAAATGTATAGTATTTATTAATAGATA 3695
431 CTTACATCAGAGATCCCAAGAAAGGAATTTTATTTAATGCAATTTGAACCATGCCCT 490
3696 CATATGTAAGAGATAATATAGAAAAAATGCAATTTATTTAACGCTATAGAAACAATGGAAT 3755
491 ATGTTAAGAAAAAGCAGATTTGGGCTTTCGATGATGATAGCAGATAGAAAAATCTACTTTTG 550
3756 GCGTAAAAAGAAAGCTGATTTGGGCGAGAAATGATATC---TAGCAACAGGTATATG 3812
551 GGGAAAGAGTGGTGGCTTTTCTGCTGTAGAAAGAGATTTTCTTCTCAGGATCTTTTGTCTG 610
3813 GAGAAAGAGTAGTAGCAATTTGCGAGCTGTGGAGGGAATATCTTTCTGCTTCAATTTGCTG 3872
611 CTATATCTGCTAAAGAAAGAGAGGTCTTATGCGAGGACTCACTTTTCCAAATGAATCA 670
3873 CTATATTTTGGATAAAAAAGCGAGATTTGATGCCGGATTAACATTTTCTTAATGAACATA 3932
671 TCAGCAGAGATGAAGGACTTCACTCTGAGCTTTGCTTGTGCTGATGTTCCAAATCTTAGTAA 730
3933 TAAGTAGAGACGAAGGTTTACATTTGTGATTTTTCGCTGTTTAAAGTTTAAACATTTATATAC 3992
731 ATAAGCCTTCAGAAAGAGGCTCAGGAGATCATTTGTTGATGCTGTCAAAATTTGACGAG 790
3993 ATCCACCATCTAAGGAAGTTTAACTCGATATCATTTGATGCGGTATATATAGAAAGG 4052
791 AGTTTTTAAAGAGGCTTTCAGGTTGCGCTCATTTGGAATGAATTTGATTTGATGAAC 850
4053 AGTTTTTGACAGTTGCTATTCGGTGGATCTTATAGGTATGAATTTGTTGTTAATGCTC 4112
851 AGTACATTTGACTTTGTAGCTGACAGATTTCTGTTGGAATTTGATTTCTCAAGGTTTTC 910
4113 AGTATATAGAATTCGTCGAGATAGATTTATTAACAGAGTTAGGTTGTGGAAG--TCTCA 4170
911 AGCGAGAAATCCTTTTTCATTTTATGGAATAACATTTTCTTTAGAAGAAAAACAAATTTCT 970
4171 ATGTATATATCTTTTATGCTTTTATGAGTATATATCACTAGAGTATAGACTTAATTTT 4230
971 TTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGATTATGCGAGAAACACAGATAACG 1030
4231 TCGAACGAGGATTTAGTCAATATCAABAGATGCGGGTGTGTTTACAAATAAGAGAGAATA 4290
1031 TCTTCACCTTGGAGTCAGATTTT 1053
4291 TATTTTCTCGGATATAGATTTT 4313

RESULT 14
US-09-299-268-1/c
; Sequence 1, Application US/09299268
; Patent No. 6217882
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi ueia, Eladio
; APPLICANT: Gibbs, R.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Foxvirus as a
; TITLE OF INVENTION: Live Vaccine Vector
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
```

STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: U.S.A.
ZIP: 32606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/09/299,268
APPLICATION NUMBER: US/09/299,268
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/901,127

FILING DATE: 08/901,127
APPLICATION NUMBER: US 07/908,241
FILING DATE: 1-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/908,630
FILING DATE: 29-JUN-1992
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/342,212
FILING DATE: 21-APR-1992
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UF35.1.FWCC1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 14176 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 3852..4226
FEATURE:
NAME/KEY: CDS
LOCATION: 4585..4887
FEATURE:
NAME/KEY: CDS
LOCATION: 5131..5310
FEATURE:
NAME/KEY: CDS
LOCATION: 5760..5912
FEATURE:
NAME/KEY: CDS
LOCATION: 6786..7130
FEATURE:
NAME/KEY: CDS
LOCATION: 10148..10513
US-09-299-268-1

Query Match 40.88; Score 429.8; DB 3; Length 14176;
Best Local Similarity 66.28; Pred. No. 5.9e-120; Indels 5; Gaps 2;
Matches 651; Conservative 0; Mismatches 327;
Qy 71 AAATGTAATAAGTCAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGCGGTTTGTCA 130
Db 10841 AAATGATTTTATACAGAAATGAGGCTATCTTCAAGAGTCTGATCTAGGTTCTTTA 10782
Qy 131 TTTTCCAAATCCAGTACCTGATTTTGGAAATGTATAAACAAGGCACAGGCTTCCTTCT 190
Db 10781 TTTTCCCTATTAAGTATCATGATATCTGGAATATGTATAAACAATCATGTCGAAGTTT 10722

Qy 191 GGCAGCAGAGAGGTCGACTTATCAAAGGATCTCCCTCACTGGAAACAAGCTTAAAGCAG 250
Db 10721 GGACCGTTGAAGAGTAGATTTATCAAAGATTTAGATGATTTGGATAAATTAACTAAG 10662
Qy 251 ATGAGAAAGTACTTCATCTCTCAGATCTTTAGCTTTTGGAGCCAGTGAATGTAA 310
Db 10661 ACCGAAATATCTTTATATAAACATATCTAGCATTTTGGCATCTAGTGTATGTAA 10602
Qy 311 ATGAAATTTTGGTGGAGCGCTTTAGTCAGGAGTGGAGTTCCAGAGCTCGCTCTTCT 370
Db 10601 ATGAGAAATTTAGCGGAAAGATTTTATGTGGATGTACAGTGTTCAGAGGCAAGATTTCT 10542
Qy 371 ATGCTTTTCAAATTTCTCATCGAGATGTTCCTCAGAGATGTACAGTTTGTGATAGACA 430
Db 10541 ATGGAATTTCAAATAGCTATGGAATATTCATTAGAAATGTATAGTTTATTAATAGATA 10482
Qy 431 CTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTTGAAACCATGCCCT 490
Db 10481 CATATGTAAGAGATAATATAGAAATAATGCAATTTTAAACGCTATAGAAACAATGGAAT 10422
Qy 491 ATGTTAAGAAAGAGCAGATTTGGGCTTGGATGTAGATAGAGATAGAAAATCTACTTTTG 550
Db 10421 GCGTAAAGAAAGAGCTGATTTGGCCAGAAATGATATC--TAGCAACAGGTATATG 10365
Qy 551 GGGAAAGAGTGGTGGCCCTTTGCTCTGTAGAGAGATTTTCTCTCAGAGATCTTTTGTCTG 610
Db 10364 GAGAAAGAGTAGTAGCATTTGACGCTGTGGAGGAATATCTTTCTGGTTCAITTTGCTG 10305
Qy 611 CTATATTTCTGGCTTAAAGAGAGAGGTCTTATGCGAGACTCACCTTTTCCAAATGAACCTCA 670
Db 10304 CTATATTTTGGATAAAGAAAGAGGATTTGATGCGCGGATTAACATTTTCTAATGAACCTAA 10245
Qy 671 TCAGCAGAGATGAAGGACTTCACCTGTGCTTGGCTTGCCTGATGTTCCCAATCTAGTAA 730
Db 10244 TAAGTAGAGAGAGGTTTACATTTGCGTGTGTTAATGTTTAAACATTTATTATAC 10185
Qy 731 ATAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCAAAATTTGAGCAGG 790
Db 10184 ATCCACCATCTAGGAAGTTTATAACGTCGATATCATTTGATCGGTTTATATAGAAAGG 10125
Qy 791 AGTTTAAACAGAGCCTTCCAGTTCGCTCATTTGGAATGAATTTGCAATTTGATGAAC 850
Db 10124 AGTTTTTGACAGTTGCTTATTCGGTGGATCTTATAGGTATGAATTTGTTTAAATGTCTC 10065
Qy 851 AGTACATTTAGTTTGTAGCTGACAGATTTCTTGGAACTTGGATTTCAAGGTTTTTC 910
Db 10064 AGTATATAGAAATTCGTCGAGATAGTTTAAACAGAGTTAGGTTGTGANAAG--TCTCA 10007
Qy 911 AGCAGAAATCTTTTGTATTTTATGAAACAATTTCTTTAGAAAGAAAAACAATTTCT 970
Db 10006 ATGTATATATCTTTTGTAGCTTTATGGATATATATCATAGAGGTAAAGACTAATTTT 9947
Qy 971 TTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACCAAGATAACG 1030
Db 9946 TCGAACGACGAGTTAGTGAATATCAAAGATGGGGTGTTTTACAAATAAGAGAGATA 9887
Qy 1031 TCTTCACCTTGGATGAGATTTT 1053
Db 9886 TATTTTCTACGGATATAGATTTT 9864

RESULT 15
US-09-299-268-14
; Sequence 14, Application US/09299268
; Patent No. 6217882
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; TITLE OF INVENTION: Live Vaccine Vector
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:

ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: U.S.A.
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/09/299,268
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/901,127
FILING DATE:
APPLICATION NUMBER: US 07/908,241
FILING DATE: 1-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/908,630
FILING DATE: 29-JUN-1992
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/342,212
FILING DATE: 21-APR-1992
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UP35.1.FWCC1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 14176 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 138..1460
FEATURE:
NAME/KEY: CDS
LOCATION: 2456..2659
FEATURE:
NAME/KEY: CDS
LOCATION: 2809..3030
FEATURE:
NAME/KEY: CDS
LOCATION: 3070..3330
FEATURE:
NAME/KEY: CDS
LOCATION: 3356..4180
FEATURE:
NAME/KEY: CDS
LOCATION: 4392..5894
FEATURE:
NAME/KEY: CDS
LOCATION: 6171..6398
FEATURE:
NAME/KEY: CDS
LOCATION: 6447..6875
FEATURE:
NAME/KEY: CDS
LOCATION: 6928..7431
FEATURE:
NAME/KEY: CDS
LOCATION: 7454..7858
FEATURE:

NAME/KEY: CDS
LOCATION: 7895..8155
FEATURE:
NAME/KEY: CDS
LOCATION: 8215..8682
FEATURE:
NAME/KEY: CDS
LOCATION: 8715..9539
FEATURE:
NAME/KEY: CDS
LOCATION: 9562..10272
FEATURE:
NAME/KEY: CDS
LOCATION: 10316..11908
FEATURE:
NAME/KEY: CDS
LOCATION: 11971..12780
FEATURE:
NAME/KEY: CDS
LOCATION: 12829..13107
FEATURE:
NAME/KEY: CDS
LOCATION: 13149..14171
US-09-299-268-14
Query Match 40.8%; Score 429.8; DB 3; Length 14176;
Best Local Similarity 66.2%; Pred. No. 5.9e-120;
Matches 651; Conservative 0; Mismatches 327; Indels 5; Gaps 2;
QY 71 AAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTCTTCGCCGGTTTGTCA 130
DB 3336 AAAATGATTTTATACACGAATGGAGCTATTTCTCAAGAGTCTGATTTCTAGGTTGTTA 3395
QY 131 TCTTTCCAAATCCAGTACCTGATATTTGGAAAAATGATAAAACAGGACAGGTTCTCTTCT 190
DB 3396 TTTTCCCTTAAAGTATCATGATATCTGAAAAATGTATAAACAAATCAGTGCGCAAGTTT 3455
QY 191 GGACAGCAGAAAGAGTGCAGCTATCAAGAGGATCTCCCTCACTGGAAACAGCTTAAGCAG 250
DB 3456 GGRCCGTTGAAGAGAGTAGATTTATCAAAAGATTTAGATGATTTGGGATAAATTAACATAAG 3515
QY 251 ATGAGAAGTACTTCTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTGTAGTAATTTGTA 310
DB 3516 ACGAAATATCTTTATAAAACATATACAGCATTTTTCATCTAGTGTAGTATTTGTA 3575
QY 311 ATGAAAAATTTGGTGGAGCGCTTTTAGTCAGAGGAGTGCAGGTTCCAGAGGCTCGCTGTTCT 370
DB 3576 ATGAGAAATTTAGCGGAAGATTTTATGTTGATGTACAGTGTTCAGAGGCCAGATGTTCT 3635
QY 371 ATGGCTTTCAAATTTCTCATCGAGAAATGTTCTACAGAGATGTACAGTTTGTGCTGATAGACA 430
DB 3636 ATGGATTTCAAATAGCTATGGAATAATTCATTCAGAAATGTATAGTTTTTATTAATAGATA 3695
QY 431 CTATACATCAGAGATCCCAAGAAAGGGAAATTTTATTTAAATGCAATTTGAAACCATGCCCT 490
DB 3696 CATATGTAAGAGATAATATAGAAAAAATGCAATTTATTTAACCTATAGAAACAATGGAAT 3755
QY 491 ATGTTAAAGAAAAAGCAGATTTGGGCTTTGCGATGTAGTACGATAGAAAAATCTACTTTTG 550
DB 3756 GCGTAAAAAAGAAAGCTGATTTGGGCCAGAAATGGAATC---TAGCAACAGGTATATG 3812
QY 551 GGGAAAGAGTGTGGCTTTGCTGCTGTAGAGAGGATTTCTTCTCAGAGATCTTTTGTGCTG 610
DB 3813 GAGAAAGAGTAGTAGCATTTTCAGCTGTGGAGGGAATATCTTTTCTGCTTCAATTTGCTG 3872
QY 611 CTATATTTCTGCTAAAGAGAGAGTCTTATGCCAGAGTCACTTTTCCCAATGAACTCA 670
DB 3873 CTATATTTTGGATAAAAAACAGGATTTGATGCCCGGATTAACATTTTCTTAATGAACTAA 3932
QY 671 TCAGCAGAGATGAAGGAGCTTCACCTGTGACTTTGCTTGTGCTGATGTTCATAATCTAGTAA 730
DB 3933 TAAGTAGAGAGAGGTTTACATTTGAGTTTGGCTGTTTAAATGTTTAAACATTTATATAC 3992

Qy	731	ATAAGCCTTCAGAAAGGTCAGGAGATCATTTGTGTGATGCTGTCAAAATTGAGCAGG	790
Db	3993	ATCCACCATCTAAGGAAGTTATAACGTCGATAATCAATTGATCGGTTAATATAGAAAAGG	4052
Qy	791	AGTTTTTAACAGAGCCTTGCAGTTGGCCTCATTTGGAATGAATTCATTTGATGAAC	850
Db	4053	AGTTTTTGACAGTTGCTTATTCGGTGGATCTTATAGGTATGAATTTGTTTAAATGTCTC	4112
Qy	851	AGTACATTTGAGTTTGTAGCTGACAGATTACTTTGTGGAACCTTGGATTCTCAAAGGTTTTTC	910
Db	4113	AGTATATAGAAATTCGTCGCAGATAGATTATTAACAGAGTTAGGTTGTGAAAAG--TCTCA	4170
Qy	911	AGGCAGAAAATCCCTTTTGAATTTATGGAACAATTTCTTTAGAAAGGAAAAACAATTTCT	970
Db	4171	ATGTATATAATCCCTTTTAGCTTTATGGAGTATATCACTAGAAGGTAAGACTAAATTTT	4230
Qy	971	TTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACCCACAGATAACG	1030
Db	4231	TCGAACCGACGAGTTAGTGAATATCAAAAGATGGGGGTGTTTACAAATAAAGAGAATA	4290
Qy	1031	TCTTCACCTTGGATGCAGATTTT	1053
Db	4291	TATTTCTACGGATATAGATTTT	4313

Search completed: September 26, 2005, 02:25:52
Job time : 233 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 26, 2005, 01:35:20 ; Search time 577 Seconds
(without alignments)
12201.951 Million cell updates/sec

Title: US-10-698-228-2

Perfect score: 1053
Sequence: 1 atgggcaccgcggaaggcc.....tcacctggatgcagatttt 1053

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 7400704 seqs, 3343079526 residues

Total number of hits satisfying chosen parameters: 14801408

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
- 20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
- 21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
- 22: /cgn2_6/ptodata/2/pubpna/US10J_NEW_PUB.seq.*
- 23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
- 24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1053	100.0	1053	17	US-10-698-228-2
2	1053	100.0	4955	18	US-10-172-118-71
3	1053	100.0	4955	18	US-10-342-887-71
4	1053	100.0	4955	18	US-10-698-228-3
5	1051.4	99.8	1053	18	US-10-698-228-12
6	1051.4	99.8	1081	18	US-10-698-228-4
7	578.4	54.9	1989	9	US-09-925-301-505

ALIGNMENTS

RESULT 1

US-10-698-228-2
; Sequence 2, Application US/10698228
; Publication No. US20040072253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-2
Query Match 100.0%; Score 1053; DB 18; Length 1053;
Best Local Similarity 100.0%; Pred. No. 8.5e-295;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 ATGGCGACCCGGAAGCCGAGCGCGCGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 ATGGCGACCCGGAAGCCGAGCGCGCGCTGGATCAGGATGAGAGATCATCTTCA 60
Qy 61 GACACCAAGAAAGTGAATAAGTCAAAATGAAGAGCCATCTCCCTAAGAAAGATTCGCG 120
Db 61 GACACCAAGAAAGTGAATAAGTCAAAATGAAGAGCCATCTCCCTAAGAAAGATTCGCG 120
Qy 121 CGGTTTCTCATCTTCCAAATCCAGTACCTGATATTTGAAATGTAACAGGCAAG 180
Db 121 CGGTTTCTCATCTTCCAAATCCAGTACCTGATATTTGAAATGTAACAGGCAAG 180
Qy 181 GCTTCTCTTCTGAGCAGCAGAGAGGTGCACTTATCAAAAGGATCTCCCTCACTGGAACAAG 240
Db 181 GCTTCTCTTCTGAGCAGCAGAGAGGTGCACTTATCAAAAGGATCTCCCTCACTGGAACAAG 240
Qy 241 CTTAAAGCAGATGAAGATCTTCATCTCTCACATCTTAGCTTTTGTGAGCCAGTGAT 300
Db 241 CTTAAAGCAGATGAAGATCTTCATCTCTCACATCTTAGCTTTTGTGAGCCAGTGAT 300
Qy 301 GGAATTCGTAATGAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAAGGCT 360
Db 301 GGAATTCGTAATGAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAAGGCT 360
Qy 361 CGCTGTTTCTATGCTTTTCAAAATCTCATCGAGATGTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGCTTTTCAAAATCTCATCGAGATGTTCACTCAGAGATGTACAGTTTG 420
Qy 421 CTGATAGACATTTACATCAGAGATGCCAAGAAAAGGGAATTTTATTAATGCAATGAA 480
Db 421 CTGATAGACATTTACATCAGAGATGCCAAGAAAAGGGAATTTTATTAATGCAATGAA 480
Qy 481 ACCATGCCCTATGTTAAGAAAAGAGCAGATTGGGCCCTTGCATCGATGATAGAGAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAAGAGCAGATTGGGCCCTTGCATCGATGATAGAGAAA 540
Qy 541 TCTACTTTTGGGAAAGAGTGGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 600
Qy 601 TCTTTTGTCTGCTATTTCTGGCTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
Db 601 TCTTTTGTCTGCTATTTCTGGCTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
Qy 661 AATGAATCTCATCAGCAGATGAAGGACTTCACGTGACTTTGCTTGCCTGATGTTCAA 720
Db 661 AATGAATCTCATCAGCAGATGAAGGACTTCACGTGACTTTGCTTGCCTGATGTTCAA 720
Qy 721 TACTTAGTAAATAGCCCTTCAGAAAGAAAGGTTCAGGAGATCATTTGTGATGCTGCAAA 780
Db 721 TACTTAGTAAATAGCCCTTCAGAAAGAAAGGTTCAGGAGATCATTTGTGATGCTGCAAA 780
Qy 781 ATTGAGCAGGAGTTTTTAAACAGAGCCCTTGCAGTTCGCTCATTTGGAATGAATTCATT 840
Db 781 ATTGAGCAGGAGTTTTTAAACAGAGCCCTTGCAGTTCGCTCATTTGGAATGAATTCATT 840
Qy 841 TTGATGAAAACGATACATGATGTTGTAGCTGACAGATTAATTGTGGAACCTTGGATTTCTCA 900
Db 841 TTGATGAAAACGATACATGATGTTGTAGCTGACAGATTAATTGTGGAACCTTGGATTTCTCA 900
Qy 901 AAGGTTTTTCAAGCAGAAAATCCTTTTGAATTTTATGGAACAATTTCTTTAGAGGAAA 960
Db 901 AAGGTTTTTCAAGCAGAAAATCCTTTTGAATTTTATGGAACAATTTCTTTAGAGGAAA 960
Qy 961 ACAAAATTTCTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTTGCAGTATGGCAGAAACC 1020
Db 961 ACAAAATTTCTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTTGCAGTATGGCAGAAACC 1020
Qy 1021 ACAGATAACGTTTCACTTGGATGCAATTTT 1053
Db 1021 ACAGATAACGTTTCACTTGGATGCAATTTT 1053
```

```
RESULT 2
US-10-172-118-71
; Sequence 71, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 71
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AB036063
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-71

Query Match 100.0%; Score 1053; DB 17; Length 4955;
Best Local Similarity 100.0%; Pred. No. 2.2e-294;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCGACCCGGAAGCCGAGCGCGCGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 245 ATGGCGACCCGGAAGCCGAGCGCGCGCTGGATCAGGATGAGAGATCATCTTCA 304
Qy 61 GACACCAAGAAAGTGAATAAGTCAAAATGAAGAGCCACTCTTAAGAAAGATTCGCG 120
Db 305 GACACCAAGAAAGTGAATAAGTCAAAATGAAGAGCCACTCTTAAGAAAGATTCGCG 364
Qy 121 CGGTTTGTATCTTTTCCAAATCCAGTACCTCATATTTGGAATAATGTAATAAAGGACACAG 180
Db 365 CGGTTTGTATCTTTTCCAAATCCAGTACCTCATATTTGGAATAATGTAATAAAGGACACAG 424
Qy 181 GCTTCTCTTCTGGAAGAGAGAGGTGACTTATCAAAAGGATCTCCCTCACTGGAACAAG 240
Db 425 GCTTCTCTTCTGGAAGAGAGAGGTGACTTATCAAAAGGATCTCCCTCACTGGAACAAG 484
Qy 241 CTTAAAGCAGATGAGAGTACTTCTCATCTCTCATCTTTAGCCTTTTTCAGCCAGTGAT 300
Db 485 CTTAAAGCAGATGAGAGTACTTCTCATCTCTCATCTTTAGCCTTTTTCAGCCAGTGAT 544
Qy 301 GGAATTTGTAATGAAAATTTGCTGGAGCGCTTTAGTCAGGAGGTGCAAGGCT 360
Db 545 GGAATTTGTAATGAAAATTTGCTGGAGCGCTTTAGTCAGGAGGTGCAAGGCT 604
Qy 361 CGCTGTTTCTATGCTTTTCAAAATTTCTCATCAGAGATGTTTCACTCAGAGATGTACAGTTTG 420
Db 605 CGCTGTTTCTATGCTTTTCAAAATTTCTCATCAGAGATGTTTCACTCAGAGATGTACAGTTTG 664
Qy 421 CTGATAGACATTTACATCAGAGATGCCAAGAAAAGGGAATTTTATTAATGCAATGAA 480
Db 665 CTGATAGACATTTACATCAGAGATGCCAAGAAAAGGGAATTTTATTAATGCAATGAA 724
Qy 481 ACCATGCCCTATGTTAAGAAAAGAGCAGATTGGGCCCTTGCAGTATGATAGAGAAA 540
Db 725 ACCATGCCCTATGTTAAGAAAAGAGCAGATTGGGCCCTTGCAGTATGATAGAGAAA 784
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTTCTGCTGTAGAGAGGATTTTCTTCTCAGGA 600
Db 785 TCTACTTTTGGGAAAGAGTGGTGGCTTTTCTGCTGTAGAGAGGATTTTCTTCTCAGGA 844
```

Qy 601 TCTTTTGTCTATATCTGCTAAAGAGAGAGGTCTTATGCCAGGACTCATTCTTCC 660
Db 845 TCTTTTGTCTATATCTGCTAAAGAGAGAGGTCTTATGCCAGGACTCATTCTTCC 904
Qy 661 AATGAACATCATCAGCAGAGATGAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA 720
Db 905 AATGAACATCATCAGCAGAGATGAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA 964
Qy 721 TACTTAGTAATAAGCCCTTCAAGAGAAAGGCTCAGGAGAGATCATTTGTGCTGTCAA 780
Db 965 TACTTAGTAATAAGCCCTTCAAGAGAAAGGCTCAGGAGAGATCATTTGTGCTGTCAA 1024
Qy 781 ATTGAGCAGGAGTTTAAACAGAACCTTGCAGTGTGCCCTCATTTGGAATGAATGCAAT 840
Db 1025 ATTGAGCAGGAGTTTAAACAGAACCTTGCAGTGTGCCCTCATTTGGAATGAATGCAAT 1084
Qy 841 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTAATCTTGTGGAACCTTGGATCTTCA 900
Db 1085 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTAATCTTGTGGAACCTTGGATCTTCA 1144
Qy 901 AAGTTTTTTCAGGCAGAGAAATCCTTTTGAATTTTATGAAACATTTCTTTAGAGGAAA 960
Db 1145 AAGTTTTTTCAGGCAGAGAAATCCTTTTGAATTTTATGAAACATTTCTTTAGAGGAAA 1204
Qy 961 ACAATTTCTTTGAGAACAGATTTTCAAGATATCAGCGTTTGTGAGTTATGGCAGAAACC 1020
Db 1205 ACAATTTCTTTGAGAACAGATTTTCAAGATATCAGCGTTTGTGAGTTATGGCAGAAACC 1264
Qy 1021 ACAGATAACGCTTTCACCTTGGATGACAGATTTT 1053
Db 1265 ACAGATAACGCTTTCACCTTGGATGACAGATTTT 1297

RESULT 3

US-10-342-887-71
; Sequence 71, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 71
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-71

Query Match 100.0%; Score 1053; DB 18; Length 4955;
Best Local Similarity 100.0%; Pred. No. 2.2e-294;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGGCGACCCGGAAGCGCGGAGCGGCTGGATCAGATGAGATCATCTTCA 60
Db 245 ATGGGCGACCCGGAAGCGCGGAGCGGCTGGATCAGATGAGATCATCTTCA 304
Qy 61 GACACCAACGAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120

Db 305 GACACCAACGAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 364
Qy 121 CGTTTGTCTATCTTCCCAATCCAGTACCTGATATTTGGAAATGTATATAACAGGCACAG 180
Db 365 CGTTTGTCTATCTTCCCAATCCAGTACCTGATATTTGGAAATGTATATAACAGGCACAG 424
Qy 181 GTTCTCTTCTGGACAGCAGAGAGGTGACATTTATCAAAAGGATCTCCCTCACTGGAAACAAG 240
Db 425 GTTCTCTTCTGGACAGCAGAGAGGTGACATTTATCAAAAGGATCTCCCTCACTGGAAACAAG 484
Qy 241 CTTAAAGCAGATGAGAGTACTTCACTCTCTCACTCTTAGCCTTTTTCAGCAGCAGTGTAT 300
Db 485 CTTAAAGCAGATGAGAGTACTTCACTCTCTCACTCTTAGCCTTTTTCAGCAGCAGTGTAT 544
Qy 301 GGAATGTGAATGAAATTTTGGTGAGCGCTTTAGTCAGGAGGTGCGAGTTCCAGAGGCT 360
Db 545 GGAATGTGAATGAAATTTTGGTGAGCGCTTTAGTCAGGAGGTGCGAGTTCCAGAGGCT 604
Qy 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCAGAAATGTTCACTCAGAGATGTACAGTTTG 420
Db 605 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAATGTTCACTCAGAGATGTACAGTTTG 664
Qy 421 CTGATAGACACTTATCATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTTGAA 480
Db 665 CTGATAGACACTTATCATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTTGAA 724
Qy 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCTTTCGATGATGATGATGATGATGAAA 540
Db 725 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCTTTCGATGATGATGATGATGAAA 784
Qy 541 TCTACTTTTGGGAAAGAGTGTGGCTTTCTGCTGTAGAGGAGTTTCTTCTCAGCA 600
Db 785 TCTACTTTTGGGAAAGAGTGTGGCTTTCTGCTGTAGAGGAGTTTCTTCTCAGCA 844
Qy 601 TCTTTTGTCTCTATATCTGCTAAAGAGAGAGGTCTTATGCCAGGACTCATTCTTTTCC 660
Db 845 TCTTTTGTCTCTATATCTGCTAAAGAGAGAGGTCTTATGCCAGGACTCATTCTTTTCC 904
Qy 661 AATGAACATCATCAGCAGAGATGAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 905 AATGAACATCATCAGCAGAGATGAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 964
Qy 721 TACTTAGTAATAAGCCCTTCAAGAGAAAGGCTCAGGAGAGATCATTTGTGCTGTCAA 780
Db 965 TACTTAGTAATAAGCCCTTCAAGAGAAAGGCTCAGGAGAGATCATTTGTGCTGTCAA 1024
Qy 781 ATTGAGCAGGAGTTTAAACAGAAAGCCCTTGCAGTGTGGCCTCATTTGGAATGAATGCAAT 840
Db 1025 ATTGAGCAGGAGTTTAAACAGAAAGCCCTTGCAGTGTGGCCTCATTTGGAATGAATGCAAT 1084
Qy 841 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTAATCTTGTGGAACCTTGGATCTTCA 900
Db 1085 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTAATCTTGTGGAACCTTGGATCTTCA 1144
Qy 901 AAGTTTTTTCAGGCAGAGAAATCCTTTTGAATTTTATGAAACATTTCTTTAGAGGAAA 960
Db 1145 AAGTTTTTTCAGGCAGAGAAATCCTTTTGAATTTTATGAAACATTTCTTTAGAGGAAA 1204
Qy 961 ACAATTTCTTTGAGAACAGATTTTCAAGATATCAGCGTTTGTGAGTTATGGCAGAAACC 1020
Db 1205 ACAATTTCTTTGAGAACAGATTTTCAAGATATCAGCGTTTGTGAGTTATGGCAGAAACC 1264
Qy 1021 ACAGATAACGCTTTCACCTTGGATGACAGATTTT 1053
Db 1265 ACAGATAACGCTTTCACCTTGGATGACAGATTTT 1297

RESULT 4
US-10-698-228-3
; Sequence 3, Application US/10698228
; Publication No. US20040072253A1
; GENERAL INFORMATION:

; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA

; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 3
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-3

Query Match 100.0%; Score 1053; DB 18; Length 4955;
Best Local Similarity 100.0%; Pred. No. 2.2e-294;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCGGACCGGAAAGCGCGGCGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 245 ATGGCGGACCGGAAAGCGCGGCGCTGGATCAGGATGAGAGATCATCTTCA 304
Qy 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTGC 120
Db 305 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTGC 364
Qy 121 CGTTTGTCTATCTTCCATCCAGTACCTGATATTTGGAAAATGTATAACAGGCACAG 180
Db 365 CGTTTGTCTATCTTCCATCCAGTACCTGATATTTGGAAAATGTATAACAGGCACAG 424
Qy 181 GCTTCTCTTGGACAGAGAGGTGACATTCAAGAGATCTCCCTCACTGGAAACAAG 240
Db 425 GCTTCTCTTGGACAGAGAGGTGACATTCAAGAGATCTCCCTCACTGGAAACAAG 484
Qy 241 CTTAAGCAGATGAGAAGTACTTCACTCTCCATCTTACATCTTTAGCTTTTGGACGCAAGTAT 300
Db 485 CTTAAGCAGATGAGAAGTACTTCACTCTCCATCTTACATCTTTAGCTTTTGGACGCAAGTAT 544
Qy 301 GGAATGTGAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGGCT 360
Db 545 GGAATGTGAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGGCT 604
Qy 361 CGCTGTTTCTATGCTTTCAAAATCTCATCGAGATGTTCACTCAGAGATGTACAGTTG 420
Db 605 CGCTGTTTCTATGCTTTCAAAATCTCATCGAGATGTTCACTCAGAGATGTACAGTTG 664
Qy 421 CTGATAGACATTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 665 CTGATAGACATTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA 724
Qy 481 ACCATGCCCTATGTAAGAAAAGAGCAGATGCGCTTGGCATGATAGCAGATAGAAAA 540
Db 725 ACCATGCCCTATGTAAGAAAAGAGCAGATGCGCTTGGCATGATAGCAGATAGAAAA 784
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGTTTCTTCTCAGGA 600
Db 785 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGTTTCTTCTCAGGA 844
Qy 601 TCTTTTGTGCTATATTCTGGCTAAAGAAAGAGGTCTTATGCCAGGACTCACTTTTTC 660
Db 845 TCTTTTGTGCTATATTCTGGCTAAAGAAAGAGGTCTTATGCCAGGACTCACTTTTTC 904
Qy 661 AATGAACCTCATCAGCAGATGAAGGACTTCACTGTGACTTTTGTCTGCTGATGTCCAA 720
Db 905 AATGAACCTCATCAGCAGATGAAGGACTTCACTGTGACTTTTGTCTGCTGATGTCCAA 964

RESULT 5

US-10-698-228-12
; Sequence 12, Application US/10698228
; Publication No. US20040072253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 12
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-12

Query Match 99.8%; Score 1051.4; DB 18; Length 1053;
Best Local Similarity 99.9%; Pred. No. 2.5e-294;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGGCGGACCGGAAAGCGCGGCGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 ATGGCGGACCGGAAAGCGCGGCGCTGGATCAGGATGAGAGATCATCTTCA 60
Qy 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTGC 120
Db 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTGC 120
Qy 121 CGTTTGTCTATCTTCCATCCAGTACCTGATATTTGGAAAATGTATAAAACAGGCACAG 180
Db 121 CGTTTGTCTATCTTCCATCCAGTACCTGATATTTGGAAAATGTATAAAACAGGCACAG 180
Qy 181 GCTTCTCTTGGACAGCAGAGAGGTGACATTCAAAAGGATCTCCCTCACTGGAAACAAG 240
Db 181 GCTTCTCTTGGACAGCAGAGAGGTGACATTCAAAAGGATCTCCCTCACTGGAAACAAG 240

QY 241 CTTAAGCAGATGAGAGTACTTCTCATCTCTCATCTTAGCCCTTTTTCAGCCAGTAT 300
Db 241 CTTAAGCAGATGAGAGTACTTCTCATCTCTCATCTTAGCCCTTTTTCAGCCAGTAT 300
QY 301 GGAAATGTAATGAAATTTGGTGGAGGCTTTAGTGCAGAGGTGCGAGTCCAGAGGCT 360
Db 301 GGAAATGTAATGAAATTTGGTGGAGGCTTTAGTGCAGAGGTGCGAGTCCAGAGGCT 360
QY 361 CGCTGTTTCTATGGCTTTCAATTTCTCATCGAAGATGTTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGGCTTTCAATTTCTCATCGAAGATGTTTCACTCAGAGATGTACAGTTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 480
QY 481 ACCATGCCCTATGTTAAGAAAGAGAGATTTGGGCTTTCAGATGATAGCAGATAGAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAGAGAGATTTGGGCTTTCAGATGATAGCAGATAGAAA 540
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGTCTGTAGAGAGGATTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGTCTGTAGAGAGGATTTTCTTCTCAGGA 600
QY 601 TCTTTTGTCTATATTTCTGGCTAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 660
Db 601 TCTTTTGTCTATATTTCTGGCTAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 660
QY 661 AATGAACACTCAGCAGAGATGAGGACTTCACTGTGACTTTTGTGCTGCTGATGTTCCAA 720
Db 661 AATGAACACTCAGCAGAGATGAGGACTTCACTGTGACTTTTGTGCTGCTGATGTTCCAA 720
QY 721 TACTTAGTAAATAGCCCTTTAAGAGAGAGGTGAGGAGATCATTTGTTGATGCTGTCAA 780
Db 721 TACTTAGTAAATAGCCCTTTAAGAGAGAGGTGAGGAGATCATTTGTTGATGCTGTCAA 780
QY 781 ATTGAGCAGAGTATTTTAAAGAGAGCTTTGCCAGTTGGCCTCATTTGGAATGATTCATT 840
Db 781 ATTGAGCAGAGTATTTTAAAGAGAGCTTTGCCAGTTGGCCTCATTTGGAATGATTCATT 840
QY 841 TTGATGAAACAGTACATTTAGTTGTAGTGCAGATTAATTTGTGGAATTTCTCA 900
Db 841 TTGATGAAACAGTACATTTAGTTGTAGTGCAGATTAATTTGTGGAATTTCTCA 900
QY 901 AAGGTTTTTTCAGGAGAGAAATCCTTTTCAATTTTATGGAACATTTCTTTAGAGGAAA 960
Db 901 AAGGTTTTTTCAGGAGAGAAATCCTTTTCAATTTTATGGAACATTTCTTTAGAGGAAA 960
QY 961 ACAATTTCTTTGAGAAACGAGTTTCAGATATCAGCGTTTTCAGATTTATGCGAGAAACC 1020
Db 961 ACAATTTCTTTGAGAAACGAGTTTCAGATATCAGCGTTTTCAGATTTATGCGAGAAACC 1020
QY 1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053
Db 1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053

RESULT 6

US-10-698-228-4
; Sequence 4, Application US/10698228
; Publication No. US20040072253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06

; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 4
; LENGTH: 1081
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-4

Query Match 99.8%; Score 1051.4; DB 18; Length 1081;
Best Local Similarity 99.9%; Pred. No. 2.5e-294;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGCAGCCCGAAGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 20 ATGGCGCAGCCCGAAGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 79
QY 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCACCTCTAAGAAAGAGTTCTCGC 120
Db 80 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCACCTCTAAGAAAGAGTTCTCGC 139
QY 121 CGGTTTGTCTATCTTTTCAATCCAGTACCCTGATATTTGGAAATGTATAAACAGCACAG 180
Db 140 CGGTTTGTCTATCTTTTCAATCCAGTACCCTGATATTTGGAAATGTATAAACAGCACAG 199
QY 181 GCTTCTTCTGGACAGCAGAGAGGTGCACTTATCAAGGATCTCCCTCAGTCGAAACAG 240
Db 200 GCTTCTTCTGGACAGCAGAGAGGTGCACTTATCAAGGATCTCCCTCAGTCGAAACAG 259
QY 241 CTTAAAGCAGATCAGAGAGTACTTCTCATCTCAGATTTAGCCTTTTTCAGCCAGTAT 300
Db 260 CTTAAAGCAGATCAGAGAGTACTTCTCATCTCAGATTTAGCCTTTTTCAGCCAGTAT 319
QY 301 GGAAATGTAATGAAATTTGGTGGAGGCTTTTAGTCAGAGGTGCGAGTTCAGAGGCT 360
Db 320 GGAAATGTAATGAAATTTGGTGGAGGCTTTTAGTCAGAGGTGCGAGTTCAGAGGCT 379
QY 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAAGATGTTTCACTCAGAGATGTACAGTTTG 420
Db 380 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAAGATGTTTCACTCAGAGATGTACAGTTTG 439
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 480
Db 440 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 499
QY 481 ACCATGCCCTATGTTAAGAAAGAGAGTGGGCTTTCAGATGATAGCAGATAGAAA 540
Db 500 ACCATGCCCTATGTTAAGAAAGAGAGTGGGCTTTCAGATGATAGCAGATAGAAA 559
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGTCTGTAGAGAGGATTTTCTTCTCAGGA 600
Db 560 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGTCTGTAGAGAGGATTTTCTTCTCAGGA 619
QY 601 TCTTTTGTCTATATTTCTGGCTAAGAGAGAGGTCTTATGCGAGACTCACTTTTTC 660
Db 620 TCTTTTGTCTATATTTCTGGCTAAGAGAGAGGTCTTATGCGAGACTCACTTTTTC 679
QY 661 AATGAACACTCAGCAGAGATGAGGACTTCACTGTGACTTTTGTGCTGCTGATGTTCCAA 720
Db 680 AATGAACACTCAGCAGAGATGAGGACTTCACTGTGACTTTTGTGCTGCTGATGTTCCAA 739
QY 721 TACTTAGTAAATAGCCCTTTAAGAGAGAGGTGAGGAGATCATTTGTTGATGCTGTCAA 780
Db 740 TACTTAGTAAATAGCCCTTTAAGAGAGAGGTGAGGAGATCATTTGTTGATGCTGTCAA 799
QY 781 ATTGAGCAGAGTATTTTAAAGAGGCTTGCAGAGTGGCCTCATTTGGAATGATTCATT 840
Db 800 ATTGAGCAGAGTATTTTAAAGAGGCTTGCAGAGTGGCCTCATTTGGAATGATTCATT 859
QY 841 TTGATGAAACAGTACATTTGATGAGTGTAGCTGAGATTAATTTGGAACCTTGAATTTCA 900

```

Db      860  TTGATGAAACAGATACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACTTGGATTCTCA  919
Qy      901  AAGGTTTTTTCAGGCAGAAAATCCTTTTGTATTTATGAAAAACATTTCTTTAGAAAGGAAAA  960
Db      920  AAGGTTTTTTCAGGCAGAAAATCCTTTTGTATTTATGAAAAACATTTCTTTAGAAAGGAAAA  979
Qy      961  ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC  1020
Db      980  ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC  1039
Qy      1021  ACAGATAACGTCCTTCACCTTGGATGCAGATTTT  1053
Db      1040  ACAGATAACGTCCTTCACCTTGGATGCAGATTTT  1072

RESULT 7
US-09-925-301-505
; Sequence 505, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIORITY APPLICATION NUMBER: PCT/US00/05982
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 505
; LENGTH: 1989
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1917)
; OTHER INFORMATION: n equals a.t.g, or c
US-09-925-301-505

```

Qy	509	ATTGGGCGCTTGGCGATGGAATAGCAGATAGAAAATCTACTTTTGGGGAAAGTGCTGCGCCT	568
Db	700	ACTGGGCGCTTGGCGCTGGATTTGGGGCAAAAGAGGCTACCTATGCTGAAACGCTGTTGTAGCCT	759
Qy	569	TTGCTGCTGTAGAGAGAGTTCCTCTTCAGGATCTTTTGCCTGTATATTTCTGCTCAAGA	628
Db	760	TTGCTGAGCTGGAGAGCAATTTCTTTTCCGGTTCTTTTGGCTCGATATCTGCTCAAGA	819
Qy	629	AGAGAGCTCTTATGCCAGGACTCACTCTTTTCCAAATGAACTCATCAGCAGAGATGAAGGAC	688
Db	820	AACGAGACTGATGCCCTGGCCTCACATTTTCTAAATGAACTTATTAGCAGAGATGAGGTT	879
Qy	689	TTCACTGTGTACTTTGCTTGGCTGATGTTCCAATPACTTATGAATAGACCTTCAGAGAAA	748
Db	880	TACACTGTGATTTTGTCTGTGCTGATGTTCAACACACCTGGTATACAAACCATCGGAGGAGA	939
Qy	749	GGGTGCGGAGAGATCATTTGTTGATGCTGTCAAAATTCAGCAGGAGTTCATTTAAACAGAGCCT	808
Db	940	GAGTATAGAGAAATTAATATCAATGCTGTTCGGATAGAACAGGAGTTCCTCCTGAGGCCT	999
Qy	809	TGCCAGTGTGCCCTCATTTGGAATGAAATTGCAATTTTGATGAAACAGTACATTTGAGTTGTAG	868
Db	1000	TGCTCTGTAAGCTCATTTGGGATGAATGTCACCTCTAATGAAGCAATACATTTGAGTTGTGG	1059
Qy	869	CTGACAGATTACTTTGTGGAACCTTGGATTCCTCAAGGTTTTTTCAGGCAGAAAATCCTTTTG	928
Db	1060	CAGACAGACTTATGCTGGAACCTGGGTTTTTAGCAAGGTTTTTCAGAGTAGAACCCTATTTG	1119
Qy	929	ATTTTATGGAACAATTTCTTTAGAGGAAAAACAATTTCTTTTGAACAACGAGTTTCAG	988
Db	1120	ACTTTATGGAGATAATTTCACTGGAGGAAGAAGCTAACTTCTTTTGAAGAGAGTAGGCG	1179
Qy	989	AGTATCAGCGCTTTTGCAGTTTATGGCGAAGAACCAACAGATAACGCTTCCTCACCTTGGATGCAG	1048
Db	1180	AGTATCAGAGGATGGGAGTGATGCTCAAGTCCACAGAGAATTCCTTTTACCTTCGATGCTG	1239
Qy	1049	ATTT 1052	
Db	1240	ACTT 1243	

```

RESULT 8
US-10-084-817-342
; Sequence 342, Application US/10084817
; Publication No. US20030119009A1
; GENERAL INFORMATION:
; APPLICANT: Susan Stuart
; APPLICANT: Jed G. Nuchtern
; APPLICANT: Sharon E. Plon
; APPLICANT: Jason M. Shohet
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
; FILE REFERENCE: PA-0046 US
; CURRENT APPLICATION NUMBER: US/10/084,817
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/270,784
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 365
; SOFTWARE: PERL Program
; SEQ ID NO 342
; LENGTH: 2216
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030119009A1 202239.1
US-10-084-817-342

```

Query Match	54.9%	Score 578.4	DB 15	Length 2216
Best Local Similarity	75.0%	Pred. No. 1.1e-156		
Matches 723	Conservative 0	Mismatches 241	Indels 0	Gaps 0
Qy	89	ATGAAGAGCCATCTCTTAAGAAAGAGTTCTCGCGGTTTGTTCATCTTTCACATCCAGTACC	148	

; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides									
; FILE REFERENCE: 21272-041									
; CURRENT APPLICATION NUMBER: US/10/220,335									
; CURRENT FILING DATE: 2002-08-28									
; PRIOR APPLICATION NUMBER: 09/664,641									
; PRIOR FILING DATE: 2000-09-19									
; PRIOR APPLICATION NUMBER: 09/616,807									
; PRIOR FILING DATE: 2000-07-14									
; PRIOR APPLICATION NUMBER: 09/597,707									
; PRIOR FILING DATE: 2000-06-17									
; PRIOR APPLICATION NUMBER: 09/577,409									
; PRIOR FILING DATE: 2000-05-18									
; PRIOR APPLICATION NUMBER: 09/515,126									
; PRIOR FILING DATE: 2000-02-28									
; NUMBER OF SEQ ID NOS: 688									
; SOFTWARE: Custom									
; SEQ ID NO 514									
; LENGTH: 2482									
; TYPE: DNA									
; ORGANISM: Homo sapiens									
US-10-220-335-514									
Query Match 54.9%; Score 578.4; DB 22; Length 2482;									
Best Local Similarity 75.0%; Pred. No. 1.1e-156;									
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;									
Qy	89	ATGAAGAGCCACCTCTAAGAAAGAGTTCTCGCCGGTTTGTCACTTTTCCAATCCAGTACC	148						
Db	2086	AGGATGAGCCGCTGTGAGAGAAACCCCGCCGCTTTGTTCATCTTCCCCCATCGAGTACC	2027						
Qy	149	CTGATATTTTGGAAAAATGTATAAACAGGACAGGCTTCTCTTCTGGACAGCAGAGAGGTGCG	208						
Db	2026	ATGATATCTGGCAGATGTATAGAAGGCGAGAGGCTTCTTTTGGACCGCCGAGAGGTTG	1967						
Qy	209	ACTTATCAAAAGGATCTCCCTCACTGGAAACAAGGTTAAAGCAGATGAAGAGTACTTCATCT	268						
Db	1966	ACCTCTCCAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT	1907						
Qy	269	CTCACATCTTAGCCCTTTTGGACGACGATGATGGAATGTAAATGAATTTGGTGGAGC	328						
Db	1906	CCCATGTTCTGGCTTTCTTTGACGAAGCGATGGCATAGTAATGAATAAACTTGGTGGAGC	1847						
Qy	329	GCTTTAGTCAGGAGGTGCAGGTTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAATTTCTCA	388						
Db	1846	GAITTAGCCAAAGTTTCAGATTACAGAAGCCGCTGTTTCTATGCTTCCAAATTTGCCA	1787						
Qy	389	TCGAGAATGTTCACTCAGAGATGTA CAGTTTGCTGATAGACACTTATCATCAGAGATCCCA	448						
Db	1786	TGAAAAACATACATTTCTGAAATGTA TAGTCTTCTTATTGACACACTTACATAAAAGATCCCA	1727						
Qy	449	AGAAAGGGAATTTTATTTAATGCAATTTGAACCAATGCCCTATGTTTAAAGAAAAGCAG	508						
Db	1726	AGAAAGGGAATTTCTCTTCAATGCCATTTGAAACGAATGCTTGTGTCAAGAAGAGGCAG	1667						
Qy	509	ATTGGGCTTTCGATGGATAGCAGATAGAAAAATCTACTTTTGGGGAAAGAGTGTGGCCCT	568						
Db	1666	ACTGGGCTTTCGCTGGAATGGGGAACAAGAGGCTTACTATGTTGAACGTTGTGTAGCCCT	1607						
Qy	569	TTGCTGCTGTAAGAGAGTTTCTTCTCAGATATCTTTTGTGCTGCTATATTTCTGCTAAAGA	628						
Db	1606	TTGCTGCAGTGAAGGCAATTTCTTTTCCGGTTTCTTTTTCGGTTCGATATTTCTGCTCAAGA	1547						
Qy	629	AGAGAGGCTTATGCCAGAGCTCACATTTTTCCAAATGAACTCATCAGCAGAGATGAGGAC	688						
Db	1546	AACGAGGACTGATGCTTCGCTCCCTACATTTTCTAAATGAACTTATTACGAGATGAGGGTT	1487						
Qy	689	TTCACTGTGACTTTGCTTGCCTGATGTTCCAAATACCTTAGTAATAATAGCCCTTCAGAGAAA	748						
Db	1486	TACACTGTGATTTTGTCTTGCCTGATGTTCAACACCTTGTAACAAACCATCGAGGAGA	1427						
Qy	749	GGGTCAAGGAGATCATGTTGTGATGCTGTCAAAATTAAGCAGGAGTTTTTAAACAGAGCCT	808						
Db	1426	GAGTAAGAGAAATAATTAATCAATGCTGTTCCGATAGAAACAGAGATTTCTCATCTGAGGCT	1367						

APPLICANT: Hyseq, Inc.

;; PRIOR FILING DATE: 2000-09-18
;; PRIOR APPLICATION NUMBER: US/60/234,052
;; PRIOR FILING DATE: 2000-09-20
;; PRIOR APPLICATION NUMBER: US/60/234,923
;; PRIOR FILING DATE: 2000-09-25
;; PRIOR APPLICATION NUMBER: US/60/235,134
;; PRIOR FILING DATE: 2000-09-25
;; PRIOR APPLICATION NUMBER: US/60/235,637
;; PRIOR FILING DATE: 2000-09-26
;; PRIOR APPLICATION NUMBER: US/60/235,638
;; PRIOR FILING DATE: 2000-09-26
;; PRIOR APPLICATION NUMBER: US/60/235,711
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: US/60/235,720
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: US/60/235,840
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: US/60/235,863
;; PRIOR FILING DATE: 2000-09-27
;; NUMBER OF SEQ ID NOS: 2276
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 1169
;; LENGTH: 2500
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-954-456-1169

Query Match 54.9%; Score 578.4; DB 9; Length 2500;

Best Local Similarity 75.0%; Pred. No. 1.1e-156; Indels 0; Gaps 0;

Mismatches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

Qy 89 ATGAAGAGCCACTCTCTAAGAAAGAGTTCTGCCGGTTTGTCATCTTTCCAAATCCAGTACC 148
Db 397 AGGATGAGCCGCTCTCAGAGAAACCCCGCCCTTTGTCATCTTTCCCAATCCAGTACC 456
Qy 149 CTGATATTGGAATGTATAACAGGCACAGGCTTCTCTCGACAGCAGAGAGGTG 208
Db 457 ATGATATCTGCGACATGTATAAGAGCGCAGAGGCTTCTTTTGGACCGCGAGAGGTG 516
Qy 209 ACTTATCAAGGATCTCCCTCACTGGAACAGCTTAAAGCAGATGAGAGTACTTCATCT 268
Db 517 ACCTCTCAAGGACATTCAGCAGTGGGAATCCCTGAAACCCGAGCAGAGATATTTATAT 576
Qy 269 CTCACATCTTAGCTTTTTCGACGACGATGCAATTTGTAATGAAATTTGCTGAGC 328
Db 577 CCCATGTTCTGCGCTTCTTTGACGACAGCGATGCAATGTAATGAAATTTGCTGAGC 636
Qy 329 GCTTTAGTCAGGAGTGCAGGTTCCAGAGGCTCGCTGTTTCTATGGCTTTTCAAAATTTCTCA 388
Db 637 GATTTAGCAAGAGTTTCAGATTACAGAGCCGCTGTTTCTATGGCTTCCAAATTTGCCA 696
Qy 389 TCGAAGATGTTTCACTCAGAGATGACAGTTTGTCTGATAGACACTTACATCAGAGATCCCA 448
Db 697 TGGAAAACATACATCTCGAAATGTATAGTCTTCTTATTGACACTTACATAAAGATCCCA 756
Qy 449 AGAAAGGGATTTTATTAATGCAATTTGAACCATGCCCTATGTTTAAGAAAACAGCAG 508
Db 757 AAGAAAGGGAATTTCTCTCAATGCCATTTGAACGATGCCCTTGTGTCAAGAAAGAGCAG 816
Qy 509 ATTGGGCTTTCGATGATAGCAGATAGAAATCTACTTTTGGGGGAAAGAGTGTGCGCT 568
Db 817 ACTGGGCTTTCGCTGATTTGGGACAAAGAGGCTACTATGTTGAACGTTGTAGCCT 876
Qy 569 TTGCTGCTGTAGAGGAGTTTCTTCTCAGGATCTTTTGTCTGATATTTCTGCTCTAAGA 628
Db 877 TTGCTGAGTGAAGGACATTTCTTTTCCGGTCTTTTGGTCTGATATTTCTGCTCTAAGA 936
Qy 629 AGAGAGTCTTATGCCAGGACTCACTTTTTCATGAACTCATCAGCAGAGATGAAGAC 688
Db 937 AACGAGGACTGATGCCCTGAGCTCACAATTTTCTAATGAACCTTATAGCAGAGATGAGGTT 996
Qy 689 TTCACTGTGACTTTGCTTGTGCTGATGTTCCAAATCTTAGTAATTAAGCCTTTCAGAGAA 748

Db 997 TACACTGTGATTTTCTGCTTGCCTGTGATGTTCAAACACCTGTAACAAAAACCAATCGAGGAGA 1056
Qy 749 GGGTCAGGAGATCAATTTGTCATGCTGTCAAAATTTGAGCAGAGTGTAAAAACAGAGCCT 808
Db 1057 GAGTAAGAGAAATAATTAATCAATGCTGTCGATAGAACAGAGTTCCTCACTGAGCCT 1116
Qy 809 TGCCAGTTGGCTCAATGGAATGAATTTGCAATTTGATGAAAACAGTACATTTGATTTG 868
Db 1117 TGCCTGTGAAGCTCAATGGGATGAATTCACCTCTAATGAAGCAATACATTTGAGTTGTGG 1176
Qy 869 CTGACAGATTACTTGTGGAATTGGAATTTCAAGGTTTTCAGGCGAGAAAATCCCTTTG 928
Db 1177 CAGACAGACTTATGCTGGAATTTTGTAGCAAGGTTTTCAGAGTAGAGAACCCATTG 1236
Qy 929 ATTTTATGAAAAATTTCTTTAGAGGAAACAAAATTTCTTTGAGAAACAGAGTTTCAG 988
Db 1237 ACTTTATGAGAAATATTTCTGGAAGGAAGACATACTTCTTTGAGAGAGAGTAGGCG 1296
Qy 989 AGTATCAGCGTTTTCAGTATGCGAGAAACCAAGATAACGCTTTCACCTTCGATGCAG 1048
Db 1297 AGTATCAGAGGATGGAGTGTCAAGTCCAAACAGAGAAATTTCTTTTACCTTGTGATG 1356
Qy 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 12

US-09-954-456-1827

; Sequence 1827, Application US/09954456

; Patent No. US20020115057A1

; GENERAL INFORMATION:

; APPLICANT: Young, Paul

; TITLE OF INVENTION: Sets

; FILE REFERENCE: 689290-76

; CURRENT APPLICATION NUMBER: US/09/954,456

; CURRENT FILING DATE: 2001-09-18

; PRIOR APPLICATION NUMBER: US/60/233,617

; PRIOR FILING DATE: 2000-09-18

; PRIOR APPLICATION NUMBER: US/60/234,052

; PRIOR FILING DATE: 2000-09-20

; PRIOR APPLICATION NUMBER: US/60/234,923

; PRIOR FILING DATE: 2000-09-25

; PRIOR APPLICATION NUMBER: US/60/235,134

; PRIOR FILING DATE: 2000-09-25

; PRIOR APPLICATION NUMBER: US/60/235,637

; PRIOR FILING DATE: 2000-09-26

; PRIOR APPLICATION NUMBER: US/60/235,638

; PRIOR FILING DATE: 2000-09-26

; PRIOR APPLICATION NUMBER: US/60/235,711

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: US/60/235,720

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: US/60/235,840

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: US/60/235,863

; PRIOR FILING DATE: 2000-09-27

; NUMBER OF SEQ ID NOS: 2276

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 1827

; LENGTH: 2500

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-954-456-1827

Query Match 54.9%; Score 578.4; DB 9; Length 2500;

Best Local Similarity 75.0%; Pred. No. 1.1e-156;

Mismatches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

Qy 89 ATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGCGGTTTGTGTCATCTTTCCAAATCCAGTACC 148

Db 397 AGGATGAGCCGCTCTCAGAGAAACCCCGCCGCTTTGTCTATCTTTCCCAATCCAGTACC 456

QY	149	CTGATATTTGGAAAAATGTTATAAACAGGCA CAGGCTTCCTTCTGGA CAGCAGAGAGGTGCG	208
Db	457	ATGATATCTGGCAGATGTTATAAGAAAGCAGAGGCTTCCTTTTGGACCGCGGAGGAGTTG	516
QY	209	ACTTATCAAGAGATCTCCCTCAGCTGGNACAAGCTTAAAGCAGATGAGAGTACTTTCATCT	268
Db	517	ACCTCTCCAAGGACATTCAGCATCTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT	576
QY	269	CTCACATCTTAGCGCTTTTTCGACGCCAGTGATGGAAATGTAAATGAAAAATTCGTGTGAGC	328
Db	577	CCCATGTTCTGGCTTTCTTTTGCAGCAAGCGATGGCATAGTAAATGAATAAACTTGGTGGAGC	636
QY	329	GCTTTAGTCAGGAGGTGCAAGTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAATTCCTCA	388
Db	637	GATTTAGCCAAGAAGTTTCAGATTACAGAAGCCGCTGTTTCTATGGCTTCCAAATTCGCCA	696
QY	389	TCGAGAAATGTTCACTCAGAGATGTACAGTTTGTCTGTATAGACACTTATCATCAGAGATCCCA	448
Db	697	TGAAAAACAATACATCTGAAATGTATAGTCTTCTTATTTGACACTTACATAAAAAGATCCCA	756
QY	449	AGAAAAAGGAAATTTTATTTTAAATGCAATTCGAAACCATGCCCTATGTTTAAAGAAAAAGCAG	508
Db	757	AAGAAAGGAAATTTCTCTTCATGGCCATTCGAAACGATGCCCTGTGTTCAGAGAGAGGCAG	816
QY	509	ATTGGGCTTTGGCATGGATAGCAGATAGAAAAATCTACTTTTGGGGAAAGAGTGGTGGCCT	568
Db	817	ACTGGGCTTTGCGCTGGATTTGGGGAACAAAGAGGCTACCTATGTTGTAAGTGTGTAGCCT	876
QY	569	TTGCTGCTGATAGAGAGATTTTCTTTCAGAGATCTTTTGTGCTGTATATCTGCTCAAAGA	628
Db	877	TTGCTGAGTGAAGGCAATTTTCTTTTCCGGTCTTTTTCGCTCGATATCTGCTCAAGA	936
QY	629	AGAGAGTCTTATGCCAGGACTCACATTTTCCAAATGAACTCATCAGCAGAGATGAAGGAC	688
Db	937	AACGAGGACTGATGCGCTGGCCTCACATTTTCTAATGAACTTATTAGCAGAGATGAGGGTT	996
QY	689	TTCACTGTCACTTTGCTTGCTGATGTTTCCAAATCTTAGTAAATTAAGCCCTTCAGAGAAA	748
Db	997	TACACTGTGATTTTGTGCTGATGTTTCAAACCTGTTTCAAACCTGTTACAAACCAATCGAGNGA	1056
QY	749	GGGTCAGGAGATCATTTGTTGATGCTGTCAAAATTTGACAGGAGTTTTTTAAACAGAGCCT	808
Db	1057	GAGTAAGAGAAATAATCAATGCTGTTGCGATAGAACAGGAGTTCCTCACTGAGGCCT	1116
QY	809	TGCCAGTTGGCCTCATTTGGAATGAAATGCAATTTTGATGAAACAGTACATTTGAGTTGTAG	868
Db	1117	TGCTCTGTAAGCTCATTTGGGATGAAATGTCACCTTAATGAAGCAATACATTTGAGTTGTGG	1176
QY	869	CTCAGAGATTAATTTGTGAACTTGGATCTCAAGAGTTTTTCAGGCAGAAAAATCCCTTTTG	928
Db	1177	CACAGACATTAATGCTGGAATCTGGTTTTAGCAAGTTTTTACAGGTAGTAAACCCATTTG	1236
QY	929	ATTTTATGAAAAATTTCTTTTAGAAGAAAAACAAATTTCTTTTGAAAAACGAGTTTCAG	988
Db	1237	ACTTTATGGAGATATTTTCACTGGAGGAAGACTAACTTTCTTTTGAAGAGAGTAGGCG	1296
QY	989	AGTATCAGCGTTTTTGCAAGTTATGGCAAAAAACCAAGATAACGCTTTCACCTTGGATCGAG	1048
Db	1297	AGTATCAGAGGATGGGAGTATGTCAAGTCCAAACAGAGAAATCTTTTACCTTGGATCGT	1356
QY	1049	ATTT 1052	
Db	1357	ACTT 1360	

RESULT 13

US-10-641-643-1370

US-10-641-643-1370
; Sequence 1370, Application US/10641643

; Sequence 1370, Application US/1
; Publication No. US20040077003A1

GENERAL INFORMATION:

APPLICANT: Cocks, Benjamin G.

COCKS, Benjamin
Susan G. Stuart

; Jeffrey J. Seilhamer
 ; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
 ; GENE EXPRESSION
 ; NUMBER OF SEQUENCES: 1508
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 ; STREET: 3174 PORTER DRIVE
 ; CITY: PALO ALTO
 ; STATE: CALIFORNIA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/641,643
 ; FILING DATE: 14-Aug-2003
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: <Unknown>
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Zeller, Karen J.
 ; REGISTRATION NUMBER: 37,071
 ; REFERENCE/DOCKET NUMBER: PA-0001 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (650) 855-0555
 ; TELEFAX: (650) 845-4166
 ; INFORMATION FOR SEQ ID NO: 1370:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2500 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: GENBANK
 ; CLONE: G36154
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 1370 :
 ; US-10-641-643-1370

Query Match	54.9%	Score	578.4	DB	18	Length	2500
Best Local Similarity	75.0%	Pred. No.	1.1e-156				
Matches	723	Conservative	0	Mismatches	241	Indels	0
Gaps	0						
QY	89	ATCAAGAGCCACCTCTAAGAAAGAGGTTCTCGCCGGTTTGTGCATCTTTTCCAAATCCAGTACC	148				
Db	397	AGGTTAGGCGCTGCTGAGAGAAACCCCGCCGCTTTGTTCATCTTCCCCTCGAGTACC	456				
QY	149	CTGATATTTGGAAAAATGTTATAACAGGCA CAGGCTTCTTCTGGA CAGCAGAAAGAGGTGG	208				
Db	457	ATGATATCTGGCAGATGTATAGAGGCGAGAGGCTTCTTTTGGACCGCCGAGGAGGTG	516				
QY	209	ACTTATCAAGGATCTCCCTCACTGGAA CAAGCTTAAGCAGATGAGTAGTCTTCATCT	268				
Db	517	ACCTCTCCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT	576				
QY	269	CTCACATCTTTAGGCTTTTTTTCAGCCAGGTGATGGAAATGTAAATGAAAAATTTGGTGGAGC	328				
Db	577	CCCATGTTCTGGCTTTCTTTTCAGCAAGCGATGGCATAGTAATGAAACTTCTGTTGGAGC	636				
QY	329	GCTTTTAGTCAGGAGGTGCAGGTTTCAGAGGCTCGCTGTTTCTATGCGCTTTCAAATTTCTCA	388				
Db	637	GATTTTAGCCAAGAAGTTCAGATTTACAGNAGCCCGCTGTTTCTATGCGCTTCCAAATTTGCCA	696				
QY	389	TCGAGAAATGTTTCACTCAGAGATGTACAGTTTGCTGATAGACACTTACATCAGAGATCCCA	448				
Db	697	TGGAACAATACATTTCTGAAATGTATAGTCTTCTTATTGACACTTACATAAAGATCCCA	756				
QY	449	AGAAAAAGGGAATTTTATTTAATGCAATTTGAAACCAATGCCCTATGTTAAGAAAAAGCAG	508				
Db	757	AAGAAAGGGAATTTCTCTCAATGCCATTTGAAACGATGCTCTGTGTCAAGAGAGGCAG	816				


```
QY 509 ATTGGGCTTGGCGATGATAGCAGATAGAAATCTACTTTTGGGAAAGAGTGGGCT 568
Db 817 ACTGGGCTTGGCGATGATGGGACAAAGAGGCTACCTATGGTGAACGTGTGTAGCCT 876
QY 569 TTGCTGCTGTAGAAGGAGTTTCTTCTCAGAGCTTTTGGCTGTATATCTTGGCTAAAGA 628
Db 877 TTGCTGAGTGAAGGAGTTTCTTCTTGGCTGTATATCTTGGCTCAAGA 936
QY 629 AGAGAGTCTTATGCCAGGACTCACTTTTCCAAATGAATCTCATCAGCAGAGATGAAGGAC 688
Db 937 AACGAGACTGATGGCTTGGCTCACAATTTCTAATGAATCTTATAGCAGATGAGGTT 996
QY 689 TTCACTGTGACTTTTGGCTTGGCTGATGTTCCAAATCTTAGTAAATAAGCCTTCAGAAGAAA 748
Db 997 TACACTGTGATTTTGGCTTGGCTGATGTTCAAAACACCTGGTACACAAACCATCGAGGAGA 1056
QY 749 GGGTCAGGAGATCATTTGCTGATGCTGTCAAAATTTAGCAGGAGTTTAAACAGAGCCT 808
Db 1057 GAGTAAGAGAAATAATATCAATGCTGTTCGGATAGAACAGGAGTTCTCACTGAGGCT 1116
QY 809 TGCCAGTTGGCCTCATTTGGAGTGAATGCAATTTGATGAAACAGTACATTCAGTTGTAG 868
Db 1117 TGCCTGTGAAGCTCATTTGGGATGAATTCGACTCTAATGAAGCAATACATTCAGTTGTGG 1176
QY 869 CTGACAGATTACTTGTGGAACTTGGATTTCTCAAGGTTTTCAGGCAGAAATCCTTTTG 928
Db 1177 CAGACAGACTTATGCTGGAACTGGTTTTCAGCAAGTTTTCAGAGTAGAACCATTTG 1236
QY 929 ATTTTATGGAACATTTCTTTTGAAGGAAACAAATTTCTTTGAGAAACGAGTTTCAG 988
Db 1237 ACTTTATGGAGATATTTCACTGGAAGGAAAGACTAATCTTTTGAAGAGAGTAGGCG 1296
QY 989 AGTATCAGGTTTTCAGTTATGCGAGAAACCAGATAGAGTTTCACTTGAATGAG 1048
Db 1297 AGTATCAGGAGTGGGAGTGATGTCAAGTCCACAGAGAAATTTCTTTTACCTTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360
```

RESULT 14

```
US-10-733-878-458
; Sequence 458, Application US/10733878
; Publication No. US20040224408A1
; GENERAL INFORMATION:
; APPLICANT: Jean-Philippe Girard
; APPLICANT: Francois Amalric
; APPLICANT: Myriam Rousseau
; APPLICANT: Thomas Clouaire
; TITLE OF INVENTION: THAP PROTEINS AS NUCLEAR RECEPTORS FOR
; TITLE OF INVENTION: CHROMOKINES AND ROLES IN TRANSCRIPTIONAL REGULATION, CELL
; FILE REFERENCE: BIOBANK.012A
; CURRENT APPLICATION NUMBER: US/10/733,878
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 60/432699
; PRIOR FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/485027
; PRIOR FILING DATE: 2003-07-03
; NUMBER OF SEQ ID NOS: 535
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 458
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-733-878-458
```

Query Match 54.9%; Score 578.4; DB 20; Length 2500;
Best Local Similarity 75.0%; Pred. No. 1.1e-156;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

```
QY 89 ATGAAGAGCCACTCTCTAAAGAAAGAGTTCTCGCGGTTTGTTCATCTTTTCCAAATCCAGTACC 148
Db 397 AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTTGTTCATCTTTCCCATCGAGTACC 456
QY 149 CTGATATTTGGAAATGTATAAAGAGGACAGGCTTCTTCTGGAACAGCAGAGAGAGTGC 208
Db 457 ATGATATCTGGCAGATGTATAAAGAGGACAGGCTTCTTCTTGGACCGCGCAGAGAGTGG 516
QY 209 ACTTATCAAAAGGATCTCCCTCACTCGAAACAAGCTTAAACAGATGAGAGATCTTCATCT 268
Db 517 ACCTCTCAAGAGCAATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTATAT 576
QY 269 CTCATCTTAGCCCTTTTTCAGCCAGTGTGAATTTGTAATGAAAAATTTTGGTGGAGC 328
Db 577 CCCATGTTCTGCTTTCTTTGAGCAAGCATAGTAAATGAATGAAACTTGGTGGAGC 636
QY 329 GCTTTAGTCAGAGGTGCAAGTTTCAGAGGCTCGCTGTTTCTATGGCTTTCAAAATCTCA 388
Db 637 GATTTAGCCAAAGAGTTTCAGATTAACAGAGCCCGCTGTTTCTATGGCTTCCAAATTGCCA 696
QY 389 TCGAGATGTTCACTCAGAGATGTACAGTTTGTCTGATACACACTTACATCAGAGATCCCA 448
Db 697 TGAAGAACATACATCTGAAATGTATAGTCTTTATTGACACTTACATATAAAGATCCCA 756
QY 449 AGAAAGGGAATTTTATTAATGCAATTTGAAACCAATGCCCTATGTTAAAGAAAAAGCAG 508
Db 757 AAGAAAGGGAATTTCTTCAATGCCATTTGAAACGATGCTTGTGTCAAGAGAGAGCAG 816
QY 509 ATTGGGCTTTCGGATGAGATAGAAATCTAATTTTGGGAAAGAGTGGTGGCCT 568
Db 817 ACTGGGCTTTCGGCTGGATTTGGGACAAAGAGGCTACCTATGTTGAACGCTGTTGTAGCCT 876
QY 569 TTGCTGCTGTAGAAGGAGTTTCTTCTCAGAGATCTTTTGTCTGCTATATTTCTGGCTAAAGA 628
Db 877 TTGCTGCAAGTGAAGGCAATTTCTTTTCCGGTTCTTTTGGCTGATATTTCTGGCTCAAGA 936
QY 629 AGAGAGGCTTTATGCCAGGACTCACTTTTCCAAATGAACTCATCAGCAGAGATGAAGGAC 688
Db 937 AACGAGGACTGATGCTGGCTTCAATTTCTAATGAATCTTATAGCAGAGATGAGGTT 996
QY 689 TTCACTGTGACTTTTGGCTTGGATTTCCAAATCTTAGTAAATAGCCTTTCAGAAAGAAA 748
Db 997 TACACTGTGATTTTGGCTTGGATTTCAAACACTTGTACACAAACCATCGAGGAGA 1056
QY 749 GGGTCAGGAGATCATTTGCTGATGCTGTCAAAATTTGAGCAGAGTTTAAACAGAGCCT 808
Db 1057 GAGTAAGAGAAATAATATCAATGCTGTTCGGATAGAAACAGGAGTTCTCTCACTGAGGCT 1116
QY 809 TGCCAGTTGGCCTCATTTGGAAATGATTCATTTGATGAAACAGTACATTCAGTTGTAG 868
Db 1117 TGCCTGTGAAGCTCATTTGGGATGAATTCGACTCTAATGAAGCAATACATTCAGTTGTGG 1176
QY 869 CTGACAGATTACTTGTGGAACTTGGATTTCTCAAGGTTTTCAGGCAGAAATCCTTTTG 928
Db 1177 CAGACAGACTTATGCTGGAACTGGTTTTCAGCAAGTTTTCAGAGTAGAACCATTTG 1236
QY 929 ATTTTATGGAACATTTCTTTTGAAGGAAACAAATTTCTTTGAGAAACGAGTTTCAG 988
Db 1237 ACTTTATGGAGATATTTCACTGGAAGGAAAGACTAATCTTTTGAAGAGAGTAGGCG 1296
QY 989 AGTATCAGGTTTTCAGTTATGCGAGAAACCAGATAGAGTTTCACTTGAATGAG 1048
Db 1297 AGTATCAGAGGATGGGAGTGATGTCAAGTCCACAGAGAAATTTCTTTTACCTTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360
```

RESULT 15
US-10-843-3751
; Sequence 3751, Application US/10843641A
; Publication No. US2005006445A1

GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals, Inc.
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; TITLE OF INVENTION: Signature Gene Sets
; FILE REFERENCE: 689290-189
; CURRENT APPLICATION NUMBER: US/10/843,641A
; CURRENT FILING DATE: 2004-05-12
; PRIOR FILING DATE: US/09/873,367
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US/09/954,531
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,832
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/964,824
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/09/967,768
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/09/968,007
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,708
; PRIOR FILING DATE: 2001-10-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 8447
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3751
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-843-641A-3751

Query Match 54.9%; Score 578.4; DB 21; Length 2500;
Best Local Similarity 75.08; Pred. No. 1.1e-156;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;
QY 89 ATGAAGGCCACTCTAAGAAAGAGTTCTCGCGGTTTGTTCATCTTTTCCAATCCAGTACC 148
DB 397 AGGATGAGCGCTGCTGAGAGAAAACCCCGCGCTTTGTTCATCTTCCCATCGAGTACC 456
QY 149 CTGATATTGGAAATGTATAACAGGCACAGGCTTCTTCTGACAGCAGAGAGTCG 208
DB 457 ATGATATCTGGCAGATGATATAAGAGGCAGAGGCTTCTTTTGGACCGCGGAGGTTG 516
QY 209 ACTTATCAAGGATCTCCCTCACTGGAAACAAGCTTAAAGCAGATGAGAAGTACTTCATCT 268
DB 517 ACCTCTCAAGGACATTCAGCAGCTGGGAAATCCCTGAAACCGAGGAGAGATATTTTATAT 576
QY 269 CTCACATCTTAGCCTTTTTCAGCCAGTGTATGAAATTTGTAATGAAATTTGGTGGAGC 328
DB 577 CCCATGTTCTGGCTTTCTTTGACGAAGCAGTGGATAGTAAATGAAATCTTGGTGGAGC 636
QY 329 GCTTTAGTCAAGGAGTCAGGTTCCAGAGGCTGCTGTTCTATGGCTTTCAAAATCTCA 388
DB 637 GATTTAGCCAAAGAGTTTCAGATACAGAAGCCGCTGTTTCTATGGCTTCCAAATTTGCCA 696
QY 389 TCGAGATGTTCTCAGAGATGTACAGATTTGCTGTATGACACATTCATCAGAGATCCCA 448
DB 697 TGAAGAACATACATCTCTGAATGTATAGTCTTCTTATGACACTTACATAAAGATCCCA 756
QY 449 AGAAAGGGAATTTTATTTAATGCAATGAAACCATGCCCTATGTTAAGAAAAAAGCAG 508
DB 757 AAGAAAGGGAATTTCTCTTCAATGCCATGAAACGATGCCCTTGTGTCAAGAGAGGCAG 816
QY 509 ATTGGGCTTTCGATGATAGCAGATAGAAAATCTACTTTTGGGAAAGAGTGGTGGCCT 568
DB 817 ACTGGGCTTTCGCTGATTTGGGACAAAGAGGCTACCTATGGTGAACGTTGTAGCCT 876
QY 569 TTGCTGCTGTAGAAGGAGTTTCTTCTCAGGATCTTTTGTGCTATATCTGGGTAAGA 628

DB 877 TTGCTGAGTGAAGGCATTTTCTTTTCGGTCTTTTTCGGTGCATATTTCTGCTCAAGA 936
QY 629 AGAGAGTCTTATGCCAGGACTCACTTTTCCAAATGAATCATCATCAGCAGAGATGAAGGAC 688
DB 937 AACGAGGACTGATGCTGGCCTCACATTTTCTAAATGAATTTATTAGCAGAGATGAGGTT 996
QY 689 TTCACTGTGACTTTGCTTGGCTGATGTTCCCAATCTTAGTAAATAAGCCTTCAGAAGAAA 748
DB 997 TACACTGTGATTTTGTCTTGGCTGATGTTCAAAACACCTGCTGTACAAACCATCGGAGGA 1056
QY 749 GGGTCAGGAGATCATTTGTTGATGCTGTCAAAATTTGAGCAGGAGTTTAAACAGAGCCT 808
DB 1057 GAGTAAGAGAAATTAATTAATCAATGCTGTTCCGATAGAACAGAGTTCCTCACTGAGGCT 1116
QY 809 TGCAGTTGGCCTCATTTGGAAATGAATTCATTTTGAATGAATAAGACAGTACATTTGATTTG 868
DB 1117 TGCCTGTGAAGCTCATTTGGGATGAATTCGACTCTAATGAAGCAATACATTTGATTTG 1176
QY 869 CTGACAGATTACTTTGTGGAATCTTCAAGGTTTTCAGGCAGAGAAATCCCTTTG 928
DB 1177 CAGACAGACTTATGCTGGAATCTGGGTTTTCAGAGGTTTTCAGAGTAGAGAACCCATTTG 1236
QY 929 ATTTTATGGAACAATTTCTTTAGAAGGAAAAACAAATTTCTTTGAGAAACCGAGTTTCAG 988
DB 1237 ACTTTATGGAGATATTTCTACTGGAGGAAAGACTAATCTTTTTCAGAGAGAGTAGGCG 1296
QY 989 AGTATCAGCGTTTTCAGTGTATGAGGAGAAACCAAGATTAAGTCTTTCACCTTTGGATG 1048
DB 1297 AGTATCAGAGGATGGGAGTGTATGTCAGTCCAACAGAGAAATCTTTTACCTTTGGATG 1356
QY 1049 ATTT 1052
DB 1357 ACTT 1360

Search completed: September 26, 2005, 03:50:46
Job time : 583 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: September 25, 2005, 23:52:05 ; Search time 2732 Seconds
(without alignments)
14671.184 Million cell updates/sec

Title: US-10-698-228-2
Perfect score: 1053
Sequence: 1 atggcgaccggaaggcc.....tcacctggatgcagatttt 1053

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: gb_est1.*
2: gb_est2.*
3: gb_hic.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gss1.*
9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1053	100.0	3292	3	CR617553 full-leng
2	1053	100.0	4748	3	AL137348 Homo sapi
3	1041	98.9	4650	3	BC042948 Homo sapi
4	1010.6	96.0	3501	3	CR627376 Homo sapi
5	945	89.7	1056	9	AY398973 Homo sapi
6	929.6	88.3	1083	1	AL547501 AL547501
7	829.8	78.8	1008	9	AY398974 Pan trogl
8	793.2	75.3	879	5	BQ441857 AGENCOURT
9	777	73.8	1043	9	AY398975 Mus muscu
10	731.2	69.4	892	7	CN163214 952695 MA
11	728.2	69.2	914	7	CF995079 AGENCOURT
12	710.6	67.5	870	7	CO648380 ILLUMIGEN
13	682.8	64.8	804	6	CD656661 AGENCOURT
14	655.4	62.2	1061	4	EM801298 AGENCOURT
15	654.8	62.2	896	5	BU170979 AGENCOURT
16	616.4	58.5	685	5	BX951374 DXFzp781E
17	578.4	54.9	1571	3	CR603461 full-leng
18	578.4	54.9	1573	3	CR625489 full-leng
19	578.4	54.9	1582	3	CR608076 full-leng
20	578.4	54.9	1588	3	CR602054 full-leng
21	578.4	54.9	1592	3	CR621427 full-leng
22	578.4	54.9	1600	3	CR604378 full-leng
23	578.4	54.9	1605	3	CR590959 full-leng
24	578.4	54.9	1612	3	CR614990 full-leng

25	578.4	54.9	1613	3	CR609838 full-leng
26	578.4	54.9	1623	3	CR618451 full-leng
27	578.4	54.9	1630	3	CR602150 full-leng
28	578.4	54.9	1796	3	CR596700 full-leng
29	570.4	54.2	2088	3	AK088907 Mus muscu
30	562	53.4	638	7	CN358022 170005326
31	547	51.9	547	7	CR416843 CR416843
32	530.8	50.4	836	7	CR416843 CR416843
33	528.8	50.2	608	1	AU137050 AU137050
34	527.2	50.1	540	5	BU429508 UI-HF-BNO
35	525	49.9	997	4	BM468712 AGENCOURT
36	518.8	49.3	522	5	BQ694828 1000971 H
37	513.6	48.8	793	7	CK778582 965745 MA
38	503.8	47.8	923	7	CR580780 CR580780
39	503	47.8	792	5	BU328458 603494277
40	502	47.7	510	7	CR540284 DXFzp459N
41	501.4	47.6	697	6	CA328968 UI-M-FY0
42	493.8	46.9	510	5	BX955728 DXFzp781G
43	493.2	46.8	946	7	CN024619 AGENCOURT
44	491.8	46.7	932	5	BU196941 AGENCOURT
45	489.4	46.5	909	6	CA981614 AGENCOURT

ALIGNMENTS

RESULT 1
CR617553
LOCUS full-length cDNA clone CS0D1011YF14 of Placentia Cot 25-normalized
DEFINITION of Homo sapiens (human).
ACCESSION CR617553
VERSION CR617553.1 GI:50498360
KEYWORDS HTC; CNSLT_CDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1. (Bases 1 to 3292)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
JOURNAL
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
Genoscope.
2. (Bases 1 to 3292)
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
1. 3292
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1011YF14"
/tissue type="Placentia Cot 25-normalized"
/plasmid="pCMVSPORT_6"

Query Match 100.0%; Score 1053; DB 3; Length 3292;
Best Local Similarity 100.0%; Pred. No. 3.3e-269;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGCGCACCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGATCATCTTCA 60
|||||
Db 27 ATGGCGCACCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGATCATCTTCA 86

Qy	61	GACACCAACGAAAGTAAATCAATGAAGAGCCACTCCCTAAGAAAGAGTTCTCGC	120
Db	87	GACACCAACGAAAGTAAATCAATGAAGAGCCACTCCCTAAGAAAGAGTTCTCGC	146
Qy	121	CGGTTTGTCACTTTCCATCCAGTACCTGATATTTGGAAATATGATAAACAAGGACAG	180
Db	147	CGGTTTGTCACTTTCCATCCAGTACCTGATATTTGGAAATATGATAAACAAGGACAG	206
Qy	181	GCTTCCTCTGGACAGAGAGGTGACATTCAAGGATCTCCCTCACTGGAACAAG	240
Db	207	GCTTCCTCTGGACAGAGAGGTGACATTCAAGGATCTCCCTCACTGGAACAAG	266
Qy	241	CTTAAGCAGATGAGAAGTACTTCTCTCTCACATCTTAGCCTTTTTCAGCCAGTGAT	300
Db	267	CTTAAGCAGATGAGAAGTACTTCTCTCTCACATCTTAGCCTTTTTCAGCCAGTGAT	326
Qy	301	GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGCT	360
Db	327	GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGCT	386
Qy	361	CGCTGTTTCTATGCTTTCAAAATTTCTCATCGAATGTTCACTCAGAGATGTACAGTTG	420
Db	387	CGCTGTTTCTATGCTTTCAAAATTTCTCATCGAATGTTCACTCAGAGATGTACAGTTG	446
Qy	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA	480
Db	447	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA	506
Qy	481	ACCATGCCCTATGTTAAGAAAAAGCAGATGGCCCTTGCAGATGATAGAGATAA	540
Db	507	ACCATGCCCTATGTTAAGAAAAAGCAGATGGCCCTTGCAGATGATAGAGATAA	566
Qy	541	TCTACTTTTGGGAAAGAGTGGTGGCCCTTGTCTCTGTAGAGAGTTTCTTCTCAGGA	600
Db	567	TCTACTTTTGGGAAAGAGTGGTGGCCCTTGTCTCTGTAGAGAGTTTCTTCTCAGGA	626
Qy	601	TCCTTTGCTGCTATATCTTGCTTAAAGAAAGAGTCTTATGCCAGGACTCACCTTTTCC	660
Db	627	TCCTTTGCTGCTATATCTTGCTTAAAGAAAGAGTCTTATGCCAGGACTCACCTTTTCC	686
Qy	661	AATGAACCTCATCAGCAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA	720
Db	687	AATGAACCTCATCAGCAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA	746
Qy	721	TACTTAGTAATAGCCTTCAGAAAGAGGTGAGGAGATCATTTGTGATGCTGTCAA	780
Db	747	TACTTAGTAATAGCCTTCAGAAAGAGGTGAGGAGATCATTTGTGATGCTGTCAA	806
Qy	781	ATTGAGCAGGAGTTTAAACAGAGCCCTTGCAGTGGCCCTCATTTGGAAATGCAATT	840
Db	807	ATTGAGCAGGAGTTTAAACAGAGCCCTTGCAGTGGCCCTCATTTGGAAATGCAATT	866
Qy	841	TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTACTTTGTGGAACCTTGGATTCTCA	900
Db	867	TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTACTTTGTGGAACCTTGGATTCTCA	926
Qy	901	AAGGTTTTTTCAGGAGAAAATCCTTTTGAATTTATGGAACAATTTCTTTAGAGAGAAA	960
Db	927	AAGGTTTTTTCAGGAGAAAATCCTTTTGAATTTATGGAACAATTTCTTTAGAGAGAAA	986
Qy	961	ACAAATTTCTTTGAGAACGAGTTTTCAGATATCAGCGTTTTCAGTTATGCGCAGAAC	1020
Db	987	ACAAATTTCTTTGAGAACGAGTTTTCAGATATCAGCGTTTTCAGTTATGCGCAGAAC	1046
Qy	1021	ACAGATAACGCTTTTCACTTGGATGCAGATTTT	1053
Db	1047	ACAGATAACGCTTTTCACTTGGATGCAGATTTT	1079
LOCUS	HSMB02033	4748 bp	mRNA linear HTC 22-SEP-2004

RESULT 2
HSMB02033

DEFINITION	Homo sapiens mRNA; cDNA DKFZp761E1312 (from clone DKFZp761E1312).
ACCESSION	AL137348
VERSION	AL137348.1 GI:6807859
KEYWORDS	HTC.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	1 (bases 1 to 4748)
AUTHORS	Ansorge, W., Krieger, S., Regiert, T., Rittmueller, C., Schwager, B., Mewes, H.W., Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and Wiemann, S.
CONSTRM	The German cDNA Consortium
TITLE	Direct Submission
JOURNAL	Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
COMMENT	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkiz-heidelberg.de; Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp761E1312) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering: http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp761E1312 Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.
FEATURES	Location/Qualifiers
source	1..4748
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="RZPD:DKFZp761E1312"
	/db_xref="taxon:9606"
	/clone="DKFZp761E1312"
	/tissue_type="amygdala"
	/clone_lib="761 (synonym: hamy2). Vector pSport1; host DH10B; sites NotI + SalI"
	/dev_stage="adult"
	/note="ribonucleotide reductase"
gene	1..4748
	/gene="DKFZp761E1312"
CDS	47..1102
	/gene="DKFZp761E1312"
	/codon_start=1
	/product="hypothetical protein"
	/protein_id="CAB70703.2"
	/db_xref="GI:52545612"
	/db_xref="GOA:Q9NTD8"
	/db_xref="UniProt/TREMBL:Q9NTD8"
	/translation="MGDPERPAAGLDQDERSSDTNESEIKSNEEPILLKSSRRPVI PPIQPDIMKMYKQASFTWAEVDLSKDLPHWNKLKADKYPFISHILAPFAASDGI VNEMLRFSQVQVPEARCFYQILLIENHSEMYSLIDITYIKDKPKRFLFNRAIE TMPYVKKADWALRWIADRSKTFGRVVAFAAVGVFFSGFPAALFWLKKRGIMPLGT FSNELISRDEGLHCDFAELQFVYNKPSSEERVREIIVDAVKIEQELTEALPVGLIG MNCILMKQYIEFVADRLIVELGFSKVFQAFNPFPFWMENISLEGKTNFPEKRVSEYQRF AVMAETTDNVFTLDAP"
ORIGIN	
Query Match	100.0%; Score 1053; DB 3; Length 4748;
Best Local Similarity	100.0%; Pred. No. 3,7e-269;
Matches 1053; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 ATGGGCGACCCGGAAGCCGGAAGCCGCGCGCTGGATCAGGATGAGATCATCTTCA 60
Db	47 ATGGGCGACCCGGAAGCCGGAAGCCGCGCGCTGGATCAGGATGAGATCATCTTCA 106
Qy	61 GACACCAAGAGTGAATAAGTCAATCAAGCCACTCCTTAAGAAAGAGTTCTCGC 120
Db	107 GACACCAAGAGTGAATAAGTCAATCAAGCCACTCCTTAAGAAAGAGTTCTCGC 166
Qy	121 CGGTTTGTCACTTTTCCAATCCAGTACCTGATATTTGGAAATATGATAAACAAGGACAG 180
Db	167 CGGTTTGTCACTTTTCCAATCCAGTACCTGATATTTGGAAATATGATAAACAAGGACAG 226

Db	72	ATGGCGGACCCGGAAGGCGCGGCTGGATCAGGATGAGAGATCATCTTCA	131
Qy	61	GACACCAAGAAAGTAAAGTCAAAATGAAGAGCCACTCCTTAAGAAAGAGTTCTCGC	120
Db	132	GACACCAAGAAAGTAAAGTCAAAATGAAGAG-CACTCCTTAAGAAAGAGTTCTCGC	190
Qy	121	CGGTTTGTCACTTTTCCAATCAGTACCTCGATATTTGGAAAAATGTATAACAGGCAAG	180
Db	191	CGGTTTGTCACTTTTCCAATCAGTACCTCGATATTTGGAAAAATGTATAACAGGCAAG	250
Qy	181	GCCTTCCTTGGACAGAGAGGTCGACTTATCAAGAGATCTCCCTCACTGGAACAA	240
Db	251	GCCTTCCTTGGACAGAGAGGTCGACTTATCAAGAGATCTCCCTCACTGGAACAA	310
Qy	241	CTTAAGCAGATGAGAGTACTTCTATCTCTCACAATCTTAGCCTTTTGGCAGCGATGAT	300
Db	311	CTTAAGCAGATGAGAGTACTTCTATCTCTCACAATCTTAGCCTTTTGGCAGCGATGAT	370
Qy	301	GGAAATGTAATGAAAAATTTGGTGGAGCGCTTTTGTAGTCAGGAGGTCAGGTTCCAGAGCT	360
Db	371	GGAAATGTAATGAAAAATTTGGTGGAGCGCTTTTGTAGTCAGGAGGTCAGGTTCCAGAGCT	430
Qy	361	CGCTGTTTCTATGCTTTCAAATTTCTCATCAGAAATGTTCACTCAGAGATGACAGTTTG	420
Db	431	CGCTGTTTCTATGCTTTCAAATTTCTCATCAGAAATGTTCACTCAGAGATGACAGTTTG	490
Qy	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAAATTTTATTTAATGCAATTGAA	480
Db	491	CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAAATTTTATTTAATGCAATTGAA	550
Qy	481	ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCTTTCGATGGATGAGATAGAAAA	540
Db	551	ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCTTTCGATGGATGAGATAGAAAA	610
Qy	541	TCTACTTTTGGGAAAGAGTGGTGGCTTTCGCTGTAGAGAGGTTTCTTCTCAGGA	600
Db	611	TCTACTTTTGGGAAAGAGTGGTGGCTTTCGCTGTAGAGAGGTTTCTTCTCAGGA	670
Qy	601	TCCTTTGCTGCTATTTCTGCTAAGAGAGAGGCTTATGCCAGGACTCACATTTTCC	660
Db	671	TCCTTTGCTGCTATTTCTGCTAAGAGAGAGGCTTATGCCAGGACTCACATTTTCC	730
Qy	720	AATGAATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTCTGCTGATGTTCCAA	720
Db	731	AATGAATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTCTGCTGATGTTCCAA	790
Qy	721	TACTTAGTAATAAGCCTTTCAGAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAAA	780
Db	791	TACTTAGTAATAAGCCTTTCAGAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAAA	850
Qy	781	ATTGAGCAGGAGTTTAAACAGAGAGCCTTGGCAGTTCCTCATTTGGAATGCAATT	840
Db	851	ATTGAGCAGGAGTTTAAACAGAGAGCCTTGGCAGTTCCTCATTTGGAATGCAATT	910
Qy	841	TTGATGAAACAGTACATTTGATGTTAGTACAGAGATTTACTTTGGAACTTGGATTTCTCA	900
Db	911	TTGATGAAACAGTACATTTGATGTTAGTACAGAGATTTACTTTGGAACTTGGATTTCTCA	970
Qy	901	AAGTTTTTTCAGGAGAAAAATCCCTTTTGAATTTTATGGAACAATTTCTTTAGAGGAAAA	960
Db	971	AAGTTTTTTCAGGAGAAAAATCCCTTTTGAATTTTATGGAACAATTTCTTTAGAGGAAAA	1030
Qy	961	ACAAATTTCTTTGAGAAACAGTTTTCAGAGATCAGCGTTTTCAGGATTTAGGAGAAACC	1020
Db	1031	ACAAATTTCTTTGAGAAACAGTTTTCAGAGATCAGCGTTTTCAGGATTTAGGAGAAACC	1090
Qy	1021	ACAGATAAGCTCTTCACTTCGATGACAGATTTT	1053
Db	1091	ACAGATAAGCTCTTCACTTCGATGACAGATTTT	1123

LOCUS	CR627376	3501 bp	mRNA	linear	HTC 22-SEP-2004
DEFINITION	Homo sapiens mRNA; cDNA DKFZp686M05248 (from clone DKFZp686M05248).				
ACCESSION	CR627376				
VERSION	CR627376.1	GI:50949847			
KEYWORDS	HTC.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 3501)				
AUTHORS	Ottenwaelder,B., Obermaier,B., Deutschenbaue,S., Schaipp,A., Mewes,H.W., Weill,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.				
CONSRM	The German cDNA Consortium				
TITLE	Direct Submission				
JOURNAL	Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY				
COMMENT	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by Medigenomix (Martinsried/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp686M05248) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering: http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp686M05248 Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.				
FEATURES	Location/Qualifiers				
source	1..3501				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/clone="DKFZp686M05248"				
	/tissue_type="salivary gland"				
	/clone_lib="686 (synonym: hlccc3). Vector pSport1_Sfi; host DH10B; sites SfiIA + SfiIB"				
	/dev_stage="adult"				
	/note="ribonucleotide reductase M2 B (TP53 inducible)"				
gene	1..3501				
	/gene="DKFZp686M05248"				
CDS	332..1225				
	/gene="DKFZp686M05248"				
	/codon_start=1				
	/product="hypothetical protein"				
	/protein_id="CAH10473.1"				
	/db_xref="GI:50949848"				
	/translation="MTKQAQASFWTAEEVDLSKDLPHNKLKADKDYFISHILAFPAASDGIIVNENLVERFSQEVQPEARCFYQIILIENHSEMSYLLIDTYIRDPKKREPLFNAIETMPYVKKADWALRWIADKSTFGERVVPAAVGVFSGSPAAIPLMKRGLMPLTFSNELISRDEGLHCDPACLMFOYLKPKPBERVRIIVDAVKIQEFTLEALPVGLIGNCILMKQYIEFVADRLVELFGSKVQAENPFPFPMENISLBGKTNPFKEKRVSEYQRFAPVMAETTDNVFTLDAF"				
ORIGIN					
Query Match	96.0%; Score 1010.6; DB 3; Length 3501;				
Best Local Similarity	98.6%; Pred. No. 6.8e-258;				
Matches 1019; Conservative	0; Mismatches 14; Indels 0; Gaps 0;				
Qy	21	GGAAAGCGCGGCTCGATCAGATGAGAGATCATCTTCAGACACCAACGAAAGTGAAT	80		
Db	190	GGACGACGCTTGAGGCGCTCAGCGCGGAGATCATCTTCAGACACCAACGAAAGTGAAT	249		
Qy	81	AAAGTCAAAATGAGAGCCACTCTTAGAAGAGTTCTCGCGGTTGTCTATCTTTCCAAT	140		
Db	250	AAAGTCAAAATGAGAGCCACTCTCTAGAAGAGTTCTCGCGGTTGTCTATCTTTCCAAT	309		
Qy	141	CCAGTACCTCGATATTTGAAAATGTATAAACAGGCACAGGCTTCTCTTGGACACAGA	200		
Db	310	CCAGTACCTCGATATTTGAAAATGTATAAACAGGCACAGGCTTCTCTTGGACACAGA	369		
Qy	201	AGAGTTCGACTTATCAAAAGGATCTCCCTCACTGGAAACAAAGCTTAAAGCAGATGAGAAGTA	260		
Db	370	AGAGTTCGACTTATCAAAAGGATCTCCCTCACTGGAAACAAAGCTTAAAGCAGATGAGAAGTA	429		


```
Qy 661 AATGAATCTATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 661 AATGAATCTATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Qy 721 TACTTAGTAATAGCCCTTCAGAGAAAGGCTCAGGAGATCATTTGATGCTGTCAAA 780
Db 721 TACTTAGTAATAGCCCTTCAGAGAAAGGCTCAGGAGATCATTTGATGCTGTCAAA 780
Qy 781 ATTGAGCAGGAGTTTAAACAGAAAGCCCTTGCCAGTTGGCCCTCATTTGCAATGCAATT 840
Db 781 ATTGAGCAGGAGTTTAAACAGAAAGCCCTTGCCAGTTGGCCCTCATTTGCAATGCAATT 840
Qy 841 TTGATGAAAACAGTACATGAGTTTGTAGCTGACAGATTACTTTGGAACCTTGGATTCTCA 900
Db 841 TTGATGAAAACAGTACATGAGTTTGTAGCTGACAGATTACTTTGGAACCTTGGATTCTCA 900
Qy 901 AAGGTTTTTTCAGGAGAAATCCTTTTGTATTTTATGAAACATTTCTTTAGAGGAAA 960
Db 901 AAGGTTTTTTCAGGAGAAATCCTTTTGTATTTTATGAAACATTTCTTTAGAGGAAA 960
Qy 961 ACAAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTTATGCGAAGAAC 1020
Db 961 ACAAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTTATGCGAAGAAC 1020
Qy 1021 ACAGATAACGTTCTTCACTTGGATGACAGATTTT 1053
Db 1021 ACAGATAACGTTCTTCACTTGGATGACAGATTTT 1053
```

RESULT 6

```
AL547501 1083 bp mRNA linear EST 25-MAR-2004
DEFINITION AL547501 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1011YF14 5-PRIME, mRNA sequence.
ACCESSION AL547501
VERSION AL547501.3 GI:45747948
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 1083)
Li, W.B., Gruber, C., Jesses, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:31269332.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4436.r
```

For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS0D1011DC07QPl&c=4436.r.

FEATURES

```
source
Location/Qualifiers
1. .1083
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1011YF14"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
```

ORIGIN

Query Match

88.3%; Score 929.6; DB 1; Length 1083;

RESULT 7

```
Best Local Similarity 98.0%; Pred. No. 1.9e-236;
Matches 976; Conservative 6; Mismatches 10; Indels 4; Gaps 4;
Qy 1 ATGGGCGACCCGGAAAGCGCGGAAAGCGCGCGG-GCTGGATCAGGATGAGAGATCATCTTTC 59
Db 27 ATGGGCGACCCGGAAAGCGCGGAAAGCGCGCGGNGCTGGATCAGGATGAGAGATCATCTTTC 86
Qy 60 AGACACCAACGAAAGTGAATAAAGTCAAAAGAGAGCGACCTCCCTAAAGAAAGAGTTCTCG 119
Db 87 AGACACCAACGAAAGTGAATAAAGTCAAAAGAGAGCGACCTCCCTAAAGAAAGAGTTCTCG 146
Qy 120 CCGGTTGTCTATCTTTTCCAATCCAGTACCCCTGATATTTGGAAAAATGATATAAAGAGCA 179
Db 147 CCGGTTGTCTATCTTTTCCAATCCAGTACCCCTGATATTTGGAAAAATGATATAAAGAGCA 206
Qy 180 GGCTTCTCTTCGGACAGCAGAGAGGTGCACTTATCAAGAGATCTCCCTCACTGGACAA 239
Db 207 GGCTTCTCTTCGGACAGCAGAGAGGTGCACTTATCAAGAGATCTCCCTCACTGGACAA 266
Qy 240 GCTTAAAGCAGATCAGAGAGTACTTCTCATCTCTCACATCTTTAGCCTTTTTCAGAGCA 299
Db 267 GCTTAAAGCAGATCAGAGAGTACTTCTCATCTCTCACATCTTTAGCCTTTTTCAGAGCA 326
Qy 300 TGGAAATGTAATGAAATTTTGGAGCGCTTTTATGTCAGGAGGTGTCAGGTTCCAGAGGC 359
Db 327 TGGAAATGTAATGAAATTTTGGAGCGCTTTTATGTCAGGAGGTGTCAGGTTCCAGAGGC 386
Qy 360 TCGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAAATGTTTCACTCAGAGATGATACGTTT 419
Db 387 TCGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAAATGTTTCACTCAGAGATGATACGTTT 446
Qy 420 GCTGATGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAAATGCAATTCGA 479
Db 447 GCTGATGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAAATGCAATTCGA 506
Qy 480 AACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTTGCGATGGATGATAGAGATAGAAA 539
Db 507 AACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTTGCGATGGATGATAGAGATAGAAA 566
Qy 540 ATCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGATTTTCTTCTCAGG 599
Db 567 ATCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGATTTTCTTCTCAGG 626
Qy 600 ATCTTTTGTGCTATATTTCTGCTTAAAGAAAGAGGCTTTTATGCGAGGAGTCACTTTTTC 659
Db 627 ATCTTTTGTGCTATATTTCTGCTTAAAGAAAGAGGCTTTTATGCGAGGAGTCACTTTTTC 686
Qy 660 CAATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCA 719
Db 687 CAATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCA 746
Qy 720 ATACTTAGTAATTAAGCCCTCAGAGAAAGGCTCAGAGAGATCATTTGATGCTGTCAA 779
Db 747 ATACTTAGTAATTAAGCCCTCAGAGAAAGGCTCAGAGAGATCATTTGATGCTGTCAA 806
Qy 780 AATTGACGAGAGTTTAAACAGAGCCCTTCCAGTTCGCCCTCATTTGGAATCAATTCAT 839
Db 807 AATTGACGAGAGTTTAAACAGAGCCCTTCCAGTTCGCCCTCATTTGGAATCAATTCAT 866
Qy 840 TTTGATGAAACAGTACATTTGATGTTTGTAGCTGACAGATTTACTTGTGGAACCTTGGATTC 898
Db 867 TTTGATGAAACAGTACATTTGATGTTTGTAGCTGACAGATTTACTTGTGGAACCTTGGATTC 926
Qy 899 CAAAGGTTTTTTCAGGAGAAAATCCTTTTGAATTTTATGAAAACATTTCTTTTGAAGGAA 958
Db 927 CAAAGGTTTTTTCAGGAGAAAATCCTTTTGAATTTTATGCG-AAAATTTCTTTTGAAGG-A 984
Qy 959 AAACAAATTTCTTTGAGAAACAGTTTTCAGAGTATC 994
Db 985 AAAMAAAATTTCTTTKAGAACAGGTTTTRAGWGCAGC 1020
```


AV398974 1008 bp DNA linear GSS 12-DEC-2003
LOCUS Pan troglodytes HCM0069 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY398974
VERSION AY398974.1 GI:39754963
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
REFERENCE
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarimal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,P., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarimal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,P., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source location/Qualifiers
1..1008
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
gene <1..>1008
/locus_tag="HCM0069"
ORIGIN
Query Match 78.8%; Score 829.8; DB 9; Length 1008;
Best Local Similarity 82.7%; Pred. No. 7.6e-210;
Matches 831; Conservative 0; Mismatches 174; Indels 0; Gaps 0;
Qy 49 AGATCATCTTCAGACACCAAGTGAATAAGTCAATGAAGCCACTCTTAAGA 108
Db 1 AGATCATCTTCAGACACCAAGTGAATAAGTCAATGAAGCCACTCTTAAGA 60
Qy 109 AAGAGTTCTCGCGGTTGTTCATCTTTCCATCCAGTACCCTGATATTTGAAATGTAT 168
Db 61 AAGAGTTCTCGCGGTTGTTCATCTTTCCATCCAGTACCCTGATATTTGAAATGTAT 120
Qy 169 AAACAGGCACAGGCTTCTTCTGACACGACGAGAGTGCATTCATCAAGAGATCTCCCT 228
Db 121 AAACAGGCACAGGCTTCTTCTGACACGACGAGAGTGCATTCATCAAGAGATCTCCCT 180
Qy 229 CACTGGACAGCTTAAGCAGATGACAGTACTTCTCATCTCTCATCTTACCTTTT 288
Db 181 CACTNGACAGCTTAAGCAGATGACAGTACTTCTCATCTCTCATCTTACCTTTT 240
Qy 289 GCAGCCAGTGTAGTGAATTTGAAATTTGTTGGAGCGCTTTAGTCAGAGGTGCAG 348
Db 241 GCAGCCAGTGTAGTGAATTTGAAATTTGTTGGAGCGCTTTAGTCAGAGGTGCAG 300
Qy 349 GTTCCAGAGGCTCGCTGTTTCTATGGCTTTTCAAAATCTCATCGAAATGTTCACTCAGAG 408
Db 301 GTTCCAGAGGCTCGCTGTTTCTATGGCTTTTCAAAATCTCATCGAAATGTTCACTCAGAG 360
Qy 409 ATGTACAGTTTGTGTAGACACTTACATCAGATCCCAAGAAAGGGAATTTTATTT 468
Db 361 ATGTACAGTTTGTGTAGACACTTACATCAGATCCCAAGAAAGGNNNNNNNNNN 420
Qy 469 AATCAATTGAACCATGCGCTTATGTTAAGAAAAAGCAGATTTGGCGCTTGGCATGATA 528

Db 421 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 480
Qy 529 GCAGATAGAAAATCTACTTTTTGGGGAAAGAGTGGTGGCCCTTTGCTGCTCTAGAGAGATT 588
Db 481 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 540
Qy 589 TTCTTCTCAGGATCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCGCAGGA 648
Db 541 TTCTTCTCAGGATCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCGNNNN 600
Qy 649 CTCACCTTTTCCAAATGAATCATCAGCAGAGATGAAGGACTTCACGTGACCTTGTCTTGC 708
Db 601 NTCACTTTTCCAAATGAATCAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 660
Qy 709 CTGATGTTCCCAATCTAGTAAATAGCCCTTCAGAAAGAGGTTCAGGAGAGATCATTTGTT 768
Db 661 CTGATGTTCCCAATCTAGTAAATAGCCCTTCAGAAAGAGGTTCAGGAGAGATCATTTGTT 720
Qy 769 GATGCTGTCAAAATTTGAGCAGAGGTTTTTAAACAGAAAGCCTTGCAGTTGGCCCTCATTTGGA 828
Db 721 GATGCTGTCAAAATTTGAGCAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 780
Qy 829 ATGAATTCATTTGATGAACAGTACATTTGATTTGTAGCTGACAGATTTCTTGTGGAA 888
Db 781 ATGAATTCATTTGATGANACAGTACATTTGATTTGTAGCTGACAGATTTCTTGTGGAA 840
Qy 889 CTGATGTTCTCAAAAGGTTTTTTCAGGACGAAATCCTTTTGTATTTATGGAACAATTTCT 948
Db 841 CTGATGTTCTCAAAAGGTTTTTTCAGGACGAAATNNNTTGTATTTATGGAANNNNNNNN 900
Qy 949 TTAGAAGGAAAAACAAATTTCTTTTTCAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGATT 1008
Db 901 TTAGAAGGAAAAACAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 960
Qy 1009 ATGCGAGAAACACAGATAAGCTTTCACCTTTGGATGACAGATTTT 1053
Db 961 ATGCGAGAAACCCAGATAAGCANNNTTCCCTTGGATGACAGATTTT 1005
RESULT 8
BQ441857 879 bp mRNA linear EST 24-MAY-2002
LOCUS AGENCOURT 7827448 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:6014179
DEFINITION 5', mRNA sequence.
ACCESSION BQ441857
VERSION BQ441857.1 GI:21180933
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 879)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM13208 row: d column: 20
High quality sequence stop: 662.
FEATURES
source location/Qualifiers
1..879
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6014179"
/tissue_type="embryonal carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_92"

/note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;

Site 2: SalI; Cloned unidirectionally; oligo-dT primed.

Average insert size 2.5 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

ORIGIN

```
Query Match      75.3%; Score 793.2; DB 5; Length 879;
Best Local Similarity 98.7%; Pred. No. 4.2e-200;
Matches 821; Conservative 0; Mismatches 8; Indels 3; Gaps 2;

Qy 1 ATGGGCGACCGGAAAGCGCGGAGCGCGGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 22 ATGGGCGACCGGAAAGCGCGGAGCGCGGGCTGGATCAGGATGAGAGATCATCTTCA 81

Qy 61 GACACCAAGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCCCTAAGAAAGAGTTCTGC 120
Db 82 GACACCAAGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCCCTAAGAAAGAGTTCTGC 141

Qy 121 CGGTTTCTCATCTTTCAATCCAGTACCTGATATTTGGAAATGTATTAACAGGCACAG 180
Db 142 CGGTTTGTCTATCTTCAATCCAGTACCTGATATTTGGAAATGTATTAACAGGCACAG 201

Qy 181 GCTTCTCTTGGACAGCAGAGAGGTGCACTTATCAAGAGTCTCCCTCACTGGAACAAG 240
Db 202 GCTTCTCTTGGACAGCAGAGAGGTGCACTTATCAAGAGTCTCCCTCACTGGAACAAG 261

Qy 241 CTAAAGCAGATGAGAGTACTTCACTCTCACATCTTAGCTTTTTCAGGCGAGTGAT 300
Db 262 CTAAAGCAGATGAGAGTACTTCACTCTCACATCTTAGGCTTTTTCAGGCGAGTGAT 321

Qy 301 GGAATTTGAATGAAATTTGGTGAGCGCTTTTAGTCAGAGGTCGAGGTTCCAGAGGCT 360
Db 322 GGAATTTGAATGAAATTTGGTGAGCGCTTTTAGTCAGAGGTCGAGGTTCCAGAGGCT 381

Qy 361 CGCTGTTTCTATGCTTTTCAAAATCTCATCAGAGATGTTTCACTCAGAGATGTACAGTTTG 420
Db 382 CGCTGTTTCTATGCTTTTCAAAATCTCATCAGAGATGTTTCACTCAGAGATGTACAGTTTG 441

Qy 421 CTGATAGACACTTACATCAGAGATGCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 442 CTGATAGACACTTACATCAGAGATGCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 501

Qy 481 ACCATGCCCTATGTTAGAAAAAGCAGATGGGCTTCGATGATGATGAGATAGAAAA 540
Db 502 ACCATGCCCTATGTTAGAAAAAGCAGATGGGCTTCGATGATGATGAGATAGAAAA 561

Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTTCTGCTGTAGAGAGATTTTCTCTCAGGA 600
Db 562 TCTACTTTTGGGAAAGAGTGGTGGCTTTTCTGCTGTAGAGAGATTTTCTCTCAGGA 621

Qy 601 TCTTTTGTCTGCTATATTTCTGCTAAAGAGAGAGTCTTATGCCAGGACTCACATTTTTC 660
Db 622 TCTTTTGTCTGCTATATTTCTGCTAAAGAGAGAGTCTTATGCCAGGACTCACATTTTTC 681

Qy 661 AATGAATCTATCAGCAGAGATGAAGACTTCACTGTGACTTTTGTGCTGATGTTCCAA 720
Db 682 AATGAATCTATCAGCAGAGATGAAGACTTCACTGTGACTTTTGTGCTGATGTTCCAA 741

Qy 721 TACTTAGTAAATAGCCCTTCAGAGAAAGGG--TCAGGGAGATCATTTGATGCTGTCA 778
Db 742 TACTTAGTAAATAGCCCTTCAGAGAAAGGGGTTCAGGGAGATCATTTGATGCTGTCA 801

Qy 779 AAATT--GAGCAGGAGTATTTTAAACAGAGCCCTTGGCAGTTGGCCCTCAATTGGA 829
Db 802 AAATTGGAGCAGGAGTATTTTAAACAGAGCCCTTGGCAGTTGGCCCTCAATTGGA 853
```

RESULT 9
AY398975
LOCUS

DEFINITION Mus musculus HCM0069 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY398975
VERSION AY398975.1 GI:39754964
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1043)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,P., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1043)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,P., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source location/Qualifiers
1..1043
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>1043
/locus_tag="HCM0069"
gene
ORIGIN
Query Match 73.8%; Score 777; DB 9; Length 1043;
Best Local Similarity 81.2%; Pred. No. 9.2e-196;
Matches 854; Conservative 0; Mismatches 185; Indels 13; Gaps 1;
Qy 1 ATGGCGCACCGGAAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 ATGGCGCACCGGAAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Qy 61 GACACCAAGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTGC 120
Db 61 GACACCAAGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTGC 120
Qy 121 CGGTTTGTCTATCTTTTCCAAATCCAGTACCTGATATTTGGAAATGTATTAACAGGCACAG 180
Db 121 CGGTTTGTCTATCTTTTCCAAATCCAGTACCTGATATTTGGAAATGTATTAACAGGCACAG 180
Qy 181 GCTTCTCTTGGACAGCAGAGAGTGCATCTTATCAAGAGTCTCCCTCACTGGAACAAG 240
Db 181 GCTTCTCTTGGACAGCAGAGAGTGCATCTTATCAAGAGTCTCCCTCACTGGAACAAG 240
Qy 241 CTAAAGCAGATGAGAGTACTTCACTCTCTCACATCTTTAGCTTTTTCAGGCGAGTGAT 300
Db 241 CTAAAGCAGATGAGAGTACTTCACTCTCTCACATCTTTAGCTTTTTCAGGCGAGTGAT 300
Qy 301 GGAATTTGAATGAAATTTGGTGAGCGCTTTTAGTCAGAGGTCGAGGTTCCAGAGGCT 360
Db 301 GGAATTTGAATGAAATTTGGTGAGCGCTTTTAGTCAGAGGTCGAGGTTCCAGAGGCT 360
Qy 361 CGCTGTTTCTATGCTTTTCAAAATCTCATCAGAGATGTTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGCTTTTCAAAATCTCATCAGAGATGTTTCACTCAGAGATGTACAGTTTG 420
Qy 421 CTGATAGACACTTACATCAGAGATGCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 421 CTAATAGACACTTACATCAGAGATGCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480

```

Qy 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATGGGCGCTTGGATGATGATAGATAGAAAA 540
Db 481 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 540
Qy 541 TCTACTTTTGGGGAAGAGTGCTGCGCTTTCTGCTGTAGAGGAGTCTTCTCTCAGGA 600
Db 541 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 600
Qy 601 TCTTTTGTCTATATTTCTGCTAAAGAGAGAGTCTTATGCCAGGACTCAGCTTTTTC 660
Db 601 TCAITTTGCTCAATATCTGCTAAGAGAGAGAGGCTCATGCTGAGTCACTTTTCA 660
Qy 661 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTGCTGATGTTCAA 720
Db 661 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTGCTGATGTTCAA 720
Qy 721 TACTTAGTAAATAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTGCTGTCAA 780
Db 721 TACTTTGTAAACAAGCCTTCAGAGATAGAGTGAGGGAATCATTTGCTGATGCTGTCAA 780
Qy 781 ATTGAGCAGGAGTATTTTAAACAGAGAGCCTTGCCAGTTGGCCCTCATTTGGAATGAA 840
Db 781 ATCAG-----CAGAGCCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 827
Qy 841 TTGATGAACAGTACATTTGATTTGTAGTGCAGAGATTTACTTTGGAACCTTTGGATTTCA 900
Db 828 TTGATGAACAGTACATTTGATTTGTAGTGCAGAGATTTACTTTGGAACCTTTGGATTTCA 887
Qy 901 AAGTTTTTTCAGGAGAAATCCTTTTGAATTTATGGAACATTTCTTTAGAGGAAAA 960
Db 888 AAGATTTTTCAGGAGAAATCCTTTTGAATTTATGGAACATTTCTTTAGAGGAAAA 947
Qy 961 ACAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020
Db 948 ACAAATTTCTTTGAGAAACAGTTTCTGAGTATCAGCGATTTTTCAGTATGGCCGAAACC 1007
Qy 1021 ACAGATAACGCTTTCACCTTGGATGCAGATTT 1052
Db 1008 ACAGATAATGCTTTCACCTTGGATGCAGATTT 1039

```

```

RESULT 10
LOCUS CN163214/c 892 bp mRNA linear EST 02-APR-2004
DEFINITION 952695 MARC 4PIG Sus scrofa cDNA 3', mRNA sequence.
ACCESSION CN163214
VERSION CN163214.1 GI:46177644
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
AUTHORS Smith,T.P.L., Freking,B.A., Ford,J.J., Vallet,J.L., Wise,T.A.,
            Noneman,D.J., Wray,J.E. and Keefe,J.W.
TITLE Porcine EST collection using a normalized library constructed from
            embryos representing early developmental stages
JOURNAL Unpublished (2003)
COMMENT Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called with phred v0.020425.c and
            trimmed with the aid of the trim_alt option. Vector identified with
            cross_match v0.990329
            Plate: TW8060 row: G column: 4
Seq primer: TAGAAGGCACAGTCGAGG.
            Location/Qualifiers
            1..892
            /organism="Sus scrofa"
            /mol_type="mRNA"
FEATURES
source

```

```

/db_xref="taxon:9823"
/tissue type="pooled"
/lab host="DH10B"
/clone lib="MARC 4PIG"
/note="Vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI;
library made with combined RNA from day-10, day-13,
day-15, day-25, and day-30 whole embryos."

ORIGIN
Query Match 69.4%; Score 731.2; DB 7; Length 892;
Best Local Similarity 94.1%; Pred. No. 1.4e-183;
Matches 760; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

Qy 246 AGCAGATCAGAAAGTACTTCTCATCTCAGATCTTAGCCCTTTTTCAGCCAGTGTGGAAT 305
Db 892 ATCAGATCAGAAAGTATTTTCTCTCATCTTAGCCCTTTTTCAGCCAGTGTGGAAT 833
Qy 306 TGTAAATGAAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGTTCAGAGGCTCGCTG 365
Db 832 TGTGAATGAAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGTTCAGAGGCTCGCTG 773
Qy 366 TTTCTATGGCTTTCAAAATCTCATCGAGATGTTTCACTCAGAGATGTACAGTTGCTGAT 425
Db 772 TTTCTATGGCTTTCAAAATCTCATCGAGATGTTTCACTCAGAGATGTACAGTTGCTGAT 713
Qy 426 AGACACTTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGGAACCAT 485
Db 712 AGACACTTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGGAACCAT 553
Qy 486 GCCCTATGTTAAGAAAAAAGCAGATTTGGGCTTTCGATGGAATGAGATAGAAAAATCTAC 545
Db 652 GCCATATGTTAAGAAAAAAGCAGATTTGGGCTTTCGATGGAATGAGATAGAAAAATCTAC 593
Qy 546 TTTTGGGAAAAAGAGTGTGGCTTTGCTGTGTAGAAAGAGTTCCTCTCAGGATCTTT 605
Db 592 TTTTGGGAAAAAGAGTGTGGCTTTGCTGTGTAGAAAGAGTTCCTCTCAGGATCTTT 533
Qy 606 TGCTGCTATATTTCTGGCTAAAGAGAGAGTTCCTATGCGAGAGTTCCTCTTTCCTCAATCA 665
Db 532 TGCTGCTATATTTCTGGCTAAAGAGAGAGTTCCTATGCGAGAGTTCCTCTTTCCTCAATCA 473
Qy 666 ACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTGCTGCTGATGTTTCCAAATCTT 725
Db 472 ACTCATCAGCAGAGATGAAGGAGTTCCTGCTGTGACTTTGCTGCTGCTGATGTTTCCAAATCTT 413
Qy 726 AGTAAATAAGCCTTTCAGAAAGAAAGGTCAGGAGATCAATTTGTGATGCTGTCAAAATCA 785
Db 412 GGTAAATAAGCCTTTCAGAAAGCAAGAGTTAGGAGAGATCAATTTGTGATGCTGTCAAAATCA 353
Qy 786 GCAGAGTTTAAACAGAGAGCCTTCCAGTTCAGGCTCATTTGGAATGAATGCAATTTGAT 845
Db 352 GCAGAGTTTAAACAGAGAGCCTTCCAGTTCAGGCTCATTTGGAATGAATGCTGTGAT 293
Qy 846 GAAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTGCGAACTTTCAGTTCCTCAAGGT 905
Db 292 GAAACAGTATATTCAGTTTGTAGCTGACAGATTAATCTGCGAACTTTCAGTTCCTCAAGGT 233
Qy 906 TTTTCAGGCGAAAAATCCTTTTGTATTTTATGGAACCAATTTCTTTAGAGGAAAAACAAA 965
Db 232 TTTTCAGGCGAAAAATCCTTTTGTATTTTATGGAACCAATTTCTTTAGAGGAAAAACAAA 173
Qy 966 TTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTCAGGAGAAACACAGA 1025
Db 172 TTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTCAGGAGAAACACAGA 113
Qy 1026 TAACGCTTTCACCTTGGATGCAGATTTT 1053
Db 112 TAATGCTTTCACCTTGGATGCAGATTTT 85

```

RESULT 11
CF995079
LOCUS

CF995079 914 bp mRNA linear EST 25-NOV-2003

DEFINITION AGENCOURT_15621478 NIH MGC 147 Homo sapiens cDNA clone
IMAGE:30528806 5', mRNA sequence.

ACCESSION CF995079

VERSION CF995079.1 GI:38511139

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 914)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgaps-remail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: NDAM615 row: f column: 15
High quality sequence stop: 677.

FEATURES
source Location/Qualifiers
1..914
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30528806"
/issue_type="Human Placenta"
/lab_host="DH10B Tona"
/clone_lib="NIH MGC 147"
/notes="Organ: placenta; Vector: pBluescriptR; Site: 1:
ali-XhoI; Site 2: BamH; Oligo-dT primed using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average
insert size 2.3 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: This is
a NIH_MGC library."

ORIGIN

Query Match 69.2%; Score 728.2; DB 7; Length 914;
Best Local Similarity 99.1%; Pred. No. 9.1e-183;
Matches 764; Conservative 0; Mismatches 3; Indels 4; Gaps 3;

QY 1 ATGGGCGACCGGAAAGCGCGGCGGCTGGATCAGATGAGATCATCTTCA 60
DB 87 ATGGGCGACCGGAAAGCGCGGCGGCTGGATCAGATGAGATCATCTTCA 146

QY 61 GACACCAACCAAGTGAATTAAGTCAATCAAGAGCCACTCTCAAGAGAGTTCTCGC 120
DB 147 GACACCAACCAAGTGAATTAAGTCAATCAAGAGCCACTCTCAAGAGAGTTCTCGC 206

QY 121 CGGTTTGTCTCTTCCATTCAGTACCTGATATTGGAAATGTATAACAGGCACAG 180
DB 207 CGGTTTGTCTCTTCCATTCAGTACCTGATATTGGAAATGTATAACAGGCACAG 266

QY 181 GCTTCCTCTTGGACAGAGAGGTGCGATTATCAAGAGTCTCCCTCACTGGAACAAG 240
DB 267 GCTTCCTCTTGGACAGAGAGGTGCGATTATCAAGAGTCTCCCTCACTGGAACAAG 326

QY 241 CTTAAAGCAGATGAGAGTACTTCACTCTCAGATCTTAGGCTTTTTCAGGCGAGTAT 300
DB 327 CTTAAAGCAGATGAGAGTACTTCACTCTCAGATCTTAGGCTTTTTCAGGCGAGTAT 386

QY 301 GGAATTGTAATGAAATTTGGTGGAGCGCTTTTAGTCAGGAGGTGCAGGTCCAGAGGCT 360

Db 387 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTCCAGAGGCT 446

QY 361 CGCTGTTTCTATGCTTTCAAATTTCTCATCGAGATGTTCACTCAGAGATGTACAGTTTG 420

Db 447 CGCTGTTTCTATGCTTTCAAATTTCTCATCGAGATGTTCACTCAGAGATGTACAGTTTG 506

QY 421 CTGATAGACACTTATCATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTGAA 480

Db 507 CTGATAGACACTTATCATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTGAA 566

QY 481 ACCATGCCCTATGTTAAG--AAAAGACAGATTGGGCTTGGCGATGATGATGATGATGAA 540

Db 567 ACCATGCCCTATGTTAAG--AAAAGACAGATTGGGCTTGGCGATGATGATGATGATGAA 624

QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTCTGCTGTAGAGAGGAGTTTCTTCTCAGGA 600

Db 625 TCTACTTTTGGGAAAGAGTGGTGGCTTCTGCTGTAGAGAGGAGTTTCTTCTCAGGA 684

QY 601 TCTTTTGCTGTATATCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 660

Db 685 TCTTTTGCTGTATATCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 744

QY 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTGTGATGTTCCAA 720

Db 745 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTGTGATGTTCCAA 804

QY 721 TACTTAGTAAATTAAGCC--TTCAGAGAAGAGGTC--AGGAGATCATTTGTTG 769

Db 805 TACTTAGTAAATTAAGCCTTTCAAAAGAAAGGCTCTAGGAAGATCATTTGTTG 855

RESULT 12

CO648380

LOCUS

DEFINITION ILLUMIGEN MCQ 41729 Katze MPB2 Macaca mulatta cDNA clone
IBUW:25149 5' similar to Bases 5 to 770 (highly similar to human
RBM2B (Hs.512592), mRNA sequence.

ACCESSION CO648380

VERSION CO648380.1 GI:50569874

KEYWORDS EST.

SOURCE Macaca mulatta (rhesus monkey)

ORGANISM Macaca mulatta

REFERENCE 1 (bases 1 to 870)

AUTHORS Katze,M.G., Thomas,M., Korth,M., Iadonato,S.P. and Magness,C.L.

TITLE Large-scale Rhesus Macaque cDNA Sequencing

JOURNAL Unpublished (2003)

COMMENT Contact: C. Magness
Illumigen Biosciences Inc.
2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
Tel: 2063780400
Fax: 2063780408
Email: cmagness@illumigen.com
Sequenced on 2004.06.02. 691 Q20 bases. Library Preparation: Prof.
Michael Katze Lab at University of Washington DNA Sequencing:
Illumigen Biosciences Inc. For further information, see
<http://www.macaque.org>

PCR Primers
FORWARD: CCCTCACTAAAGGGAACAAA
BACKWARD: CACTATAGGCGAATTGGTA
Insert Length: 870 Std Error: 0.00
Plate: CL000348 row: B column: 06
Seq primer: CCCTCACTAAAGGGAACAAA
POLYA=No.

FEATURES
source Location/Qualifiers
1..870
/organism="Macaca mulatta"
/mol_type="mRNA"
/strain="Indian"
/db_xref="taxon:9544"

[illegible]

RESULT 13	
CD656661	
LOCUS	
DEFINITION	

```
Qy 212 TATCAAGAGATCTCCCTCACTGGAAACAAGCTTAAAGCAGATGAGAAGTACTTCACTCTC 271
Db 1 TATCAAGAGATCTCCCTCACTGGAAACAAGCTTAAAGCAGATGAGAAGTACTTCACTCTC 60
Qy 272 ACATCTTAGC-CTTTTTCAGCCAGTGAATGTAATGAAATTTGGTGGAGGC 330
Db 61 ACATCTTAGCCTTTTTCAGCCAGTGAATGTAATGAAATTTGGTGGAGGC 120
Qy 331 TTTAGTCAGGAGTGCAGGTTTCAGAGCTCGCTGTTTCTATGGCTTTCAAATTTCTCATC 390
Db 121 TTTAGTCAGGAGTGCAGGTTTCAGAGCTCGCTGTTTCTATGGCTTTCAAATTTCTCATC 180
Qy 391 GAGATGTTCACTCAGAGATGTACAGTTTGTCTGTATAGACACTTACATCAGAGATCCCAAG 450
Db 181 GAGATGTTCACTCAGAGATGTACAGTTTGTCTGTATAGACACTTACATCAGAGATCCCAAG 240
Qy 451 AAAAGGGAATTTTATTTAATGCAATTAACCAATGCAATGCAATGCAATGCAATGCAATG 510
Db 241 AAAAGGGAATTTTATTTAATGCAATTAACCAATGCAATGCAATGCAATGCAATGCAATG 300
Qy 511 TGGCCCTTGGATGATAGAGATAGAAATCTACTTTTGGGGAAGAGTGTGCGCTTT 570
Db 301 TGGCCCTTGGATGATAGAGATAGAAATCTACTTTTGGGGAAGAGTGTGCGCTTT 360
Qy 571 GCTGCTGTAGAGAGATTTTCTTCTCAGGATCTTTTGTCTGTATATTTCTGCTTAAAGAAG 630
Db 361 GCTGCTGTAGAGAGATTTTCTTCTCAGGATCTTTTGTCTGTATATTTCTGCTTAAAGAAG 420
Qy 631 AGAGTCTTATGCGAGACTCACTTTTCCATGCACTCATCAGCAGAGATGAAGGACTT 690
Db 421 AGAGTCTTATGCGAGACTCACTTTTCCATGCACTCATCAGCAGAGATGAAGGACTT 480
Qy 691 CACTGTGACTTGTCTGCTGATGCTTCCAACTTCTGATGCTTCTGATGCTTCTGATGCTT 750
Db 481 CACTGTGACTTGTCTGCTGATGCTTCCAACTTCTGATGCTTCTGATGCTTCTGATGCTT 540
Qy 751 GTCAAGGAGATCAATGTTGATGCTGTCAAAATTCAGCAGAGATTTTAAACAGAGCCCTG 810
Db 541 GTCAAGGAGATCAATGTTGATGCTGTCAAAATTCAGCAGAGATTTTAAACAGAGCCCTG 600
Qy 811 CCAGTGGCCCTCATGGAATGAATGCAATTTGATGAAACAGTACATGATGATTTGTAGCT 870
Db 601 CCAGTGGCCCTCATGGAATGAATGCAATTTGATGAAACAGTACATGATGATTTGTAGCT 660
Qy 871 GACAGATCTGTTGGAACTTGGATTTCAAGGTTTTCAGGCAAGAAATCCTTTTGAT 930
Db 661 GACAGATCTGTTGGAACTTGGAT--CTCAAGGTTTTCAGGCAAGAAATCCTTTTGAT 718
Qy 931 TTTATGGAACAATTT 946
Db 719 TATGGAACAATTTT 734
```

```
RESULT 14
BM801298
LOCUS BM801298
DEFINITION AGENCOURT_6422860 NIH_MGC_92 Homo sapiens cdna clone IMAGE:5586693
5', mRNA sequence.
ACCESSION BM801298
VERSION BM801298.1 GI:19118121
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1061)
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
CONTACT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
```

cdna Library Preparation: Life Technologies, Inc.
cdna Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNL at:
<http://image.llnl.gov>
Plate: LLAM12354 row: p column: 22
High quality sequence stop: 631.

FEATURES

source

1..1061

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5586693"

/tissue_type="embryonal carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_92"

/note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;

Site 2: SalI; Cloned unidirectionally; oligo-dT primed.

Average insert size 2.5 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

ORIGIN

Query Match 62.2%; Score 655.4; DB 4; Length 1061;
Best Local Similarity 97.6%; Pred. No. 2.4e-163;
Matches 718; Conservative 0; Mismatches 12; Indels 6; Gaps 5;

```
Qy 1 ATGGGGCCACCGGAAAGCCCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 85 ATGGGGCCACCGGAAAGCCCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 144
Qy 61 GACACCAACGAAGTGAATAAAGTCAANTCAAGAGCCACTCTCTAAGAAAGAGTTCTCG 120
Db 145 GACACCAACGAAGTGAATAAAGTCAANTCAAGAGCCACTCTCTAAGAAAGAGTTCTCG 204
Qy 121 CGGTTTGTCTATCTTTCCAAATCCAGTACCTCTGATATTTGGAAATGTATAACAGGCACAG 180
Db 205 CGGTTTGTCTATCTTTCCAAATCCAGTACCTCTGATATTTGGAAATGTATAACAGGCACAG 264
Qy 181 GCTTCCTTCTGGACAGCAGAGAGTGCATTTATCAAAAGGATCTCCCTCATCTGGAACAG 240
Db 265 GCTTCCTTCTGGACAGCAGAGAGTGCATTTATCAAAAGGATCTCCCTCATCTGGAACAG 324
Qy 241 CTTAAAGCAGATGAGAAGTACTTCTATCTCTCACATCTTAGCCTTTTTCAGCCAGTGTAT 300
Db 325 CTTAAAGCAGATGAGAAGTACTTCTATCTCTCACATCTTAGCCTTTTTCAGCCAGTGTAT 384
Qy 301 GGAATTGTAAATGAAATTTTGGTGGAGCGCTTTAGTCAGAGAGTGCAGGTTCCAGAGCT 360
Db 385 GGAATTGTAAATGAAATTTTGGTGGAGCGCTTTAGTCAGAGAGTGCAGGTTCCAGAGCT 444
Qy 361 CGCTGTTTCTATGCTTTCAAATTTCTATCAGAAATGTTCTCATCAGATGTACAGTTTG 420
Db 445 CGCTGTTTCTATGCTTTCAAATTTCTATCAGAAATGTTCTCATCAGAGATGTACAGTTTG 504
Qy 421 CTGTATAGACATTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAAATGCAATTCGA 480
Db 505 CTGTATAGACATTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAAATGCAATTCGA 564
Qy 481 ACCATGCGCTTATGTTTAAAGAAAAAGCAGATTTGGGCTTTGCGATGATGATAGCAGATAGAAA 540
Db 565 ACCATGCGCTTATGTTTAAAGAAAAAGCAGATTTGGGCTTTGCGATGATGATAGCAGATAGAAA 624
Qy 541 TCTACTTTTGGGAAAAAGTGGTGGCTTTCTGCTGTGTAGAAAGGAGTTTCTTCTCAGGA 600
Db 625 TCTACTTTTGGGAAAAAGTGGTGGCTTTCTGCTGTGTAGAAAGGAGTTTCTTCTCAGGA 684
Qy 601 TCTTTTGTGCTATATTTCTGCTTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
Db 685 TCTTTTGTGCTATATTTCTGCTTAAAGAGAGAGGTC-TATGCCAGGACTCACTTTTTC 743
Qy 661 AATGAATCTCAT-CAGCAGAGAT-GAAGGACTTCACTGTGA--CTTTGCTTGCCTCAT-GT 715
```

```

Db      744 CATGACTCATCCACAGAGATGAGGAGCTTCCTGTGAACCTTGCCTTGCCTGTAGTGT 803
Qy      716 TCCAATATCTTAGTAA 731
Db      804 TCCAATATCTNAGAA 819

RESULT 15
BUI70979
LOCUS   BUI70979
DEFINITION BUI70979 896 bp mRNA linear EST 04-SEP-2002
          AGENCOURT 7940338 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6143754
          5', mRNA sequence.
BUI70979
VERSION BUI70979.1 GI:22684963
KEYWORDS EST.
SOURCE  Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 896)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgabs-x@mail.nih.gov
          Tissue Procurement: ATCC
          cDNA Library Preparation: Life Technologies, Inc.
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA sequencing by: Agencourt Bioscience Corporation
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLAM13467 row: k column: 19
          High quality sequence stop: 619.

FEATURES             source
     1..896
         Location/Qualifiers
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="IMAGE:6143754"
             /tissue_type="retinoblastoma"
             /lab_host="DH10B (phage-resistant)"
             /clone_lib="NIH_MGC_67"
             /note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
             Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
             Average insert size 1.75 kb. Library constructed by Life
             Technologies."

ORIGIN
Query Match      62.2%; Score 654.8; DB 5; Length 896;
Best Local Similarity 98.7%; Pred. No. 3.4e-163;
Matches 671; Conservative 0; Mismatches 7; Indels 2; Gaps 1;

Qy      1 ATGGGGACCCGGAAGCGCGGCTGGATCAGGATGAGATCATCTTCA 60
Db      55 ATGGGGACCCGGAAGCGCGGCTGGATGAGATCATCTTCA 114
Qy      61 GACACCAAGAAAGTGAATTAAGTCAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db      115 GACACCAAGAAAGTGAATTAAGTCAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 174
Qy      121 CGGTTTGTCTATCTTCCATCCAGTACCTCGATATTTGGAATGTATATAACAGGCACAG 180
Db      175 CGGTTTGTCTATCTTCCATCCAGTACCTCGATATTTGGAATGTATATAACAGGCACAG 234
Qy      181 GCTTCTCTTCTGGACAGAGAGGTGCGACTTATCAAGGATCTCCCTCACTGGAACAAG 240
Db      235 GCTTCTCTTCTGGACAGAGAGGTGCGACTTATCAAGGATCTCCCTCACTGGAACAAG 294
Qy      241 CTTAAAGCAGATGAGAAGTACTTTCATCTCTCAATCTTAGCCTTTTTCAGGCAGTGAT 300
Db      295 CTTAAAGCAGATGAGAAGTACTTTCATCTCTCAATCTTAGCCTTTTTCAGGCAGTGAT 354

```

```

Qy      301 GGAATTGTAAATGAAAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCT 360
Db      355 GGAATTGTAAATGAAAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCT 414
Qy      361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG 420
Db      415 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG 474
Qy      421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db      475 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 534
Qy      481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTGGGCTTTCGGATGGATAGCAGATAGAAA 540
Db      535 ACCATGCCCTATGTTAAGAAAAAAGCAGATTGGGCTTTCGGATGGATAGCAGATAGAAA 594
Qy      541 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGTCTGTAGAGGAGTTTCTTCTCAGGA 600
Db      595 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGTCTGTAGAGGAGTTTCTTCTCAGGA 654
Qy      601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 658
Db      655 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 714
Qy      659 CCAATGAACCTCATCAGCAGA 678
Db      715 CAATGAACCTCATCAGCAGA 734

```

Search completed: September 26, 2005, 02:21:58
Job time : 2741 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 1, 2005, 12:09:09 ; Search time 0.001 Seconds
(without alignments)
1478.412 Million cell updates/sec

Title: US-10-698-228-1

Perfect score: 1821

Sequence: 1 MGDPERPEAGLDQDERSSS.....QRFVMAETDNNVFLDADF 351

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2 seqs, 2106 residues

Total number of hits satisfying chosen parameters: 4

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=soft -O=seq1.pep -DB=seq -SUFFIX=pto
-OUT=swope228.res -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1
-END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALLIGN=200
-THR SCORE=ptc -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -NCPU=6 -NO XLPXY
-NEG SCORES=0 -LONGIOG -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOF=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: *seq:*

1: /home/tport/1005/swope228/align/seq2.seq.*
2: /home/tport/1005/swope228/align/seq12.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1821	100.0	1053	1	US-10-698-228-2
2	1821	100.0	1053	2	US-10-698-228-12
3	41.5	2.3	1053	1	US-10-698-228-2
4	38.5	2.1	1053	2	US-10-698-228-12

ALIGNMENTS

RESULT 1

US-10-698-228-2

; Sequence 2, Application US/10698228
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28

; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2

; LENGTH: 1053

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION:

US-10-698-228-2

Alignment Scores:

Pred. No.:	Score:	Length:
0	1821.00	1053
Percent Similarity:	100.00%	Matches: 351
Best Local Similarity:	100.00%	Conservative: 0
Query Match:	100.00%	Mismatches: 0
DB:	1	Indels: 0
		Gaps: 0

US-10-698-228-1 (1-351) x US-10-698-228-2 (1-1053)

Qy	1	MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer	20
Db	1	ATGGCGCACC CGGAAAGCGCGCGGCTGGATCAGGATGAGAGATCATCTTCA	60
Qy	21	AspThrAsnGluSerGluLeuLysSerAsnGluProLeuLeuArgLysSerArg	40
Db	61	GACACCAACGAAAGTGAATAAAGTCAATGAAGCCACTCTTAAGAAAGAGTTCGCG	120
Qy	41	ArgPheValIlePheProIleGluTy-ProAspIleTyrLysMetTyrLysGlnAlaGln	60
Db	121	CGGTTTGTCTATCTTCCAATCAGTACCTCATATTGGAAATATGATAAAGGACAG	180
Qy	61	AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys	80
Db	181	GCTTCCTCTCGACAGCAGAGAGAGGTGCATTTATCAAGAGATCTCCCTCAGTGGACAAG	240
Qy	81	LeuLysAlaAspGluLysTyrPheIleSerHisLeuAlaPheAlaAlaSerAsp	100
Db	241	CTTAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTGAT	300
Qy	101	GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla	120
Db	301	GGAAATTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTCCAGGTTCCAGAGGCT	360
Qy	121	ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu	140
Db	361	CGCTGTTTCTATGGCTTTTCAAAATTTCTATCAGAAATGTTTCACTCAGAGATGTACAGTTTG	420
Qy	141	LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu	160
Db	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTATCAATCAATTGAA	480
Qy	161	ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys	180
Db	481	ACCATGCCCTATGTTAAGAAAAAAGCAGATTGGCGCTTTCGATGGATGACAGATGAAAA	540
Qy	181	SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGluValPhePheSerGly	200
Db	541	TCTACTTTTGGGAAAGAGTGGTGGCTTTCTCTCTGTAGAGAGGAGTTTCTTCTCAGGA	600
Qy	201	SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer	220
Db	601	TCITTTGCTGCTATTTCTGGCTAAGAGAGAGAGTCTTATGCCAGGACTCACTTTTTC	660
Qy	221	AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln	240
Db	661	AATGAACATCATCAGCAGATGAAGGACTTCACCTGTGACTTGTGCTTGCTGTGCTGTCCAA	720

QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleValAlaValValLys 260
DB 721 TACTTAGTAAATAAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGTCATCTGTCAA 780
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
DB 781 ATTGACGAGGAGTTTTTAACAGAGCCTTGCAGTTGGCCTCATTTGGAATGATTTGCATT 840
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
DB 841 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTACTTTGTGGAATTCGATTCCTCA 900
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
DB 901 AAGGTTTTTCAGGCAGAAAATCCTTTGATTTTATGGAATAACATTTCTTTAGAGGAAAA 960
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
DB 961 ACAAAATTTCTTTGAGAACGAGTTTCAGAGTATCAGCGTTTTCAGATTATGCGAGAAACC 1020
QY 341 ThrAsnValPheThrLeuAspAlaAspPhe 351
DB 1021 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1053

RESULT 2

US-10-698-228-12
; Sequence 12, Application US/10698228
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA

; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; PRIOR FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14

; SEQ ID NO 12
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

; OTHER INFORMATION:

US-10-698-228-12

Alignment Scores:

Pred. No.: 0 Length: 1053
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-10-698-228-1 (1-351) x US-10-698-228-12 (1-1053)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
DB 1 ATGGGCGACCCGGAAGGCGGAGCGGCGGGCTGGATCAGATGAGATCATTTCTCA 60
QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
DB 61 GACACCAACGAAGTGAATAAAGTCAATGAAGTGAAGCCACTCTTAAGAAAGAGTTCTCGC 120
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
DB 121 CGGTTTGTTCATCTTTCCATCCAGTACCTGATTTGGAAATGATTAACAGGACAG 180
QY 61 AlaSerPheThrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80

DB 181 GCTTCTCTTGGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCCTGGAACAAG 240
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp 100
DB 241 CTTAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTAGCCTTTTTCAGCAGATGAT 300
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 361 CGCTGTTTCTATGGCTTTCAAAATTCATCAGAAATGTTTCACTCAGAGATGATCAGTTTG 420
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 421 CTGATAGACATTCATCATCAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 480
QY 161 ThrMetProTyrValLysLysAlaAspTTPalaLeuArgTTPileAlaAspArgLys 180
DB 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTGGATGCATGATAGCAGATAGAAA 540
QY 181 SerThrPheGlyGluArgValAlaPheAlaValGluGlyValPhePheSerGly 200
DB 541 TCTACTTTTGGGAAAGAGTGGGCTTTGCTGCTAGAAAGAGTTTCTTCTCAGGA 600
QY 201 SerPheAlaAlaIlePheTTPLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 601 TCTTTTGTGCTATATTTCTGGCTTAAGAAAGAGGCTTATGCCAGGACTCATTCTTCC 660
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
DB 661 AATGAACCTCATCAGCAGATGAAGGACTTCACTGTGACTTGTGCTTGTGCTGATGTTCAA 720
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAlaValLys 260
DB 721 TACTTAGTAAATAAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTCATGTCATCAA 780
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
DB 781 ATTGACGAGGAGTTTTTAACAGAGCCTTGCAGTTGGCCTCATTTGGAATGATTTGCATT 840
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
DB 841 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTACTTTGTGGAACCTTGGATTTCA 900
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
DB 901 AAGGTTTTTCAGGCAGAAAATCCTTTTGAATTTATGGAATAACATTTCTTTAGAGGAAAA 960
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
DB 961 ACAAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGATTATGCGAGAAACC 1020
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB 1021 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1053

RESULT 3

US-10-698-228-2/c
; Sequence 2, Application US/10698228
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391


```
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-2
Alignment Scores:
Pred. No.: 0 Length: 1053
Score: 41.50 Matches: 19
Percent Similarity: 47.17% Conservative: 6
Best Local Similarity: 35.85% Mismatches: 23
Query Match: 2.28% Indels: 6
DB: 1 Gaps: 3
US-10-698-228-1 (1-351) x US-10-698-228-2 (1-1053)
QY 77 HisTrp---AsnLysLeuLysAlaAspGluLysTyPheIleSerHisIleLeuAlaPhe 95
Db 298 CACTGCTGCAAAAAGGCTAAGATGTGAGATGAAGTACTTCTCATCTGCTT---TAA 242
QY 96 PheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluVal 115
Db 241 GCTTGTTCAGTGAGGGA-----GATCCTTTGATAGTGCACC-TCTTCTGCTGTC 192
QY 116 GlnValProGluAlaArgCysPheTyGlyPheGlnIle 128
Db 191 CAGAAGGAAGCCGTGCTGTTTATATACATTTTCCAAATA 153
```

RESULT 4

```
US-10-698-228-12/c
; Sequence 12, Application US/10698228
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; PRIOR FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 12
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-12
```

```
Alignment Scores:
Pred. No.: 0 Length: 1053
Score: 38.50 Matches: 30
Percent Similarity: 36.17% Conservative: 21
Best Local Similarity: 21.28% Mismatches: 63
Query Match: 2.11% Indels: 27
DB: 2 Gaps: 6
```

US-10-698-228-1 (1-351) x US-10-698-228-12 (1-1053)

```
QY 198 PheSerGlySerPheAla-----AlaIlePheTrpLeuLysLysArgGlyLeu 213
Db 977 TTCTCAAGAAATTTGTTTTTCTTAAAGAAATGTTTCCATAAAATCAAAGGATTT 918
```

```
QY 214 MetProGlyLeuThrPheSerAsn-----GluLeuIleSerArgAspGluGly 229
Db 917 TCTGCCTGAAAAACCTTTGAGAAATCCAAAGTTCCACAGTAATCTGTCACTACAACTCA 858
QY 230 LeuHisCysAspPheAlaCysLeuMetPheGlnTyLeuValAsnLysProSerGluGlu 249
Db 857 ATGTACTGT-----TTTCATCAAAATGCAATTCATTCCAATGAGGCCAACTGGCAAG 807
QY 250 ArgValArgGluIleIleValAspAlaValLysIleGluGlnGluPheLeuThrGluAla 269
Db 806 GCT-----TCTGTTAAAAAATCCTCTGCTCAATTTTGACAGCATCA 768
QY 270 LeuProValGlyLeuIleGlyMetAsnCysIleLeuMetLysGlnTyIleGluPheVal 289
Db 767 ACAATGATCTCCCTGACCCCTTTCTTCTGAAGGCTTATTACTAAGTATTGGAACATCAGG 708
QY 290 -----AlaAspArgLeuLeuValGluLeuGlyPheSerLysValPhe 303
Db 707 CAAGCAAAAGTCACAGTGAAGTCCTTCTCATCTCTGCTGATGAGTTTATGAAAAAGTGAGT 648
QY 304 GlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeu-----GluGlyLys 320
Db 647 CCTGGCATAAGACCTTCTTCTTTAGCCGAAATATAGCAGCAAAAGATCCTTGAGAGAAA 588
QY 321 Thr 321
Db 587 ACT 585
```

Search completed: November 1, 2005, 12:09:10

Job time : 1 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2005, 06:35:16 ; Search time 74 Seconds
(without alignments)
1834.499 Million cell updates/sec

Title: US-10-698-228-1

Perfect score: 1821

Sequence: 1 MGDPERPEAAGLDQDERSSS.....QRFAVMAETTDNVFTLDADF 351

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: geneeqp1980s:*
- 2: geneeqp1990s:*
- 3: geneeqp2000s:*
- 4: geneeqp2001s:*
- 5: geneeqp2002s:*
- 6: geneeqp2003as:*
- 7: geneeqp2003bs:*
- 8: geneeqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1821	100.0	351	3	AAY84439 Amino acid
2	1821	100.0	351	4	AAB69050 Human rib
3	1817	99.8	351	4	AAB93497 Human pro
4	1454	79.8	389	4	AAB69051 Human R2
5	1454	79.8	389	6	ABU07433 Protein d
6	1454	79.8	389	7	ADJ68753 Human hea
7	1454	79.8	389	8	ADJ66564 Ribonucle
8	1454	79.8	389	8	ADK70435 Respirato
9	1454	79.8	389	8	ADN03789 Antipsori
10	1454	79.8	389	8	ADN04444 Antipsori
11	1454	79.8	389	8	ADN019226 Human PRO
12	1454	79.8	389	8	ADQ09272 Human RRM
13	1454	79.8	389	8	ABM80258 Tumour-as
14	1454	79.8	413	3	AAB43302 Human can
15	1454	79.8	453	4	AAB28017 Human con
16	1454	79.8	453	4	ABG15287 Novel hum
17	1429	78.5	390	2	AAW41765 Human rib
18	1425.5	78.3	386	8	ADQ057309 Kidney de
19	1348	74.0	393	4	ABE64130 Drosophil
20	1242	68.2	430	8	ADN21469 Bacteri
21	1230	67.5	405	6	ABJ26187 Aspergill
22	1214	66.7	381	8	ADN23033 Bacteri
23	1192.5	65.5	381	6	ABJ25587 Aspergill
24	1192	65.5	413	5	ABP73700 Candida a
25	1176.5	64.6	321	8	AUS21332 Bacteri

26	1168.5	64.2	399	6	ABR53196	Abt53196 Protein s
27	1168.5	64.2	399	7	ADK63240	Adk63240 Disease t
28	1168.5	64.2	399	8	ADS43740	Ads43740 Bacterial
29	1161.5	63.8	391	8	ADS44135	Ads44135 Bacterial
30	1161.5	63.8	401	7	ADB70101	Adb70101 C. neofo
31	1130	62.1	329	3	AAG51528	Aag51528 Arabidops
32	1130	62.1	329	3	AAG05699	Aag05699 Arabidops
33	1130	62.1	332	3	AAG51527	Aag51527 Arabidops
34	1130	62.1	332	3	AAG05698	Aag05698 Arabidops
35	1130	62.1	332	3	AAG05697	Aag05697 Arabidops
36	1117	61.3	347	3	AAG14546	Aag14546 Arabidops
37	1112.5	61.1	330	3	AAG14548	Aag14548 Arabidops
38	1112.5	61.1	333	3	AAG14547	Aag14547 Arabidops
39	1099	60.4	314	3	AAG51529	Aag51529 Arabidops
40	1091	59.9	341	3	AAG41395	Aag41395 Arabidops
41	1091	59.9	363	3	AAG41394	Aag41394 Arabidops
42	1084.5	59.6	330	3	AAG41396	Aag41396 Arabidops
43	1072.5	58.9	274	2	AAW26418	Aaw26418 Swinepox
44	1072.5	58.9	274	4	AAB68239	Aab68239 Protein e
45	977.5	53.7	322	2	AAR72736	Aar72736 Plasmodiu

ALIGNMENTS

RESULT 1		
AAY84439		
ID	AAY84439 standard; protein; 351 AA.	
XX	AC	
XX	AAY84439;	
XX	AC	
DT	25-JUL-2000 (first entry)	
XX	Amino acid sequence of a human RNA-associated protein.	
DE	Human; RNA-associated protein; cell proliferation; cancer; immune response; reproductive disorder; actinic keratosis; atherosclerosis; arteriosclerosis; bursitis; cirrhosis; mixed connective tissue disease; myelofibrosis; primary paroxysmal nocturnal hemoglobinuria; polycythemia vera; trauma.	
KW	XX	
KW	XX	
KW	Homo sapiens.	
OS	Homo sapiens.	
XX	XX	
FH	Location/Qualifiers	
Key	18	
Modified-site	/note= "potential phosphorylation site"	
FT	23	
Modified-site	/note= "potential glycosylation site"	
FT	29	
Modified-site	/note= "potential phosphorylation site"	
FT	38	
Modified-site	/note= "potential phosphorylation site"	
FT	39	
Modified-site	/note= "potential phosphorylation site"	
FT	65	
Modified-site	/note= "potential phosphorylation site"	
FT	87	
Modified-site	/note= "potential phosphorylation site"	
FT	182	
Modified-site	/note= "potential phosphorylation site"	
FT	225	
Modified-site	/note= "potential phosphorylation site"	
FT	314	
Modified-site	/note= "potential glycosylation site"	
FT	329	
Modified-site	/note= "potential phosphorylation site"	
XX	XX	
XX	WO200015799-A2.	
PN	23-MAR-2000.	
XX	PD	
XX	17-SEP-1999;	
XX	99WO-US021688.	
PF	XX	

XX 17-SEP-1998; 98US-00156039.
PR 22-SEP-1998; 98US-00158720.
PR 04-NOV-1998; 98US-00186815.
PR 08-APR-1999; 99US-0128660P.
XX (INCY-) INCYTE PHARM INC.
XX Tang YT, Corley NC, Guegler KJ, Gorgone GA, Patterson C;
PI Hillman JL, Baughn MR, Lal P, Azimzai Y, Yue H, Yang J;
XX WPI; 2000-271437/23.
DR N-PSDB; AAA12411.
XX New polypeptides and polynucleotides, useful for preventing and treating
PT a disorder associated with increased or decreased expression of RNA
PT associated proteins.
XX Claim 1; Page 103-104; 131pp; English.
XX The present sequence represents a human RNA-associated protein. The
CC expression of RNA-associated proteins is closely associated with
CC reproductive tissues, nervous tissues, cell proliferation including
CC cancer, inflammation and immune responses, and so they may be used for
CC diagnosis, treatment or prevention of cell proliferative,
CC immune/inflammatory disorders, and reproductive disorders. Diseases and
CC disorders which may be treated include actinic keratosis,
CC atherosclerosis, arteriosclerosis, bursitis, cirrhosis, hepatitis, mixed
CC connective tissue disease, myelofibrosis, paroxysmal nocturnal
CC hemoglobinuria, polythermia vera, psoriasis, primary thrombocythemia
CC and cancers, and trauma
XX SQ Sequence 351 AA;
Query Match 100.0%; Score 1821; DB 3; Length 351;
Best Local Similarity 100.0%; Pred. No. 5.2e-179;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGDPERPAAGLDQDERSSTNESEIKSNEEPLLRKSSRRFVFIQYDPDIWKYKQAO 60
Db 1 MGDPERPAAGLDQDERSSTNESEIKSNEEPLLRKSSRRFVFIQYDPDIWKYKQAO 60
QY 61 ASFWTAEVDLSKDLPHWNKLKADSKYFISHILAFPAASDGI VNVNLFVERFSQEVQVPEA 120
Db 61 ASFWTAEVDLSKDLPHWNKLKADSKYFISHILAFPAASDGI VNVNLFVERFSQEVQVPEA 120
QY 121 RCFYGFQI LIENVHSEMYSLIIDTYIRDPKKREFL FNAIETMPYVKKKADWALRIADRK 180
Db 121 RCFYGFQI LIENVHSEMYSLIIDTYIRDPKKREFL FNAIETMPYVKKKADWALRIADRK 180
QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGIMPGLTFSNELISRDEGLHCDPACLMFQ 240
Db 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGIMPGLTFSNELISRDEGLHCDPACLMFQ 240
QY 241 YLVNKPSSERVREIIVDAVKIEQEFTEALPVGLIGMNCILMKQYIEFVADRLVLELGS 300
Db 241 YLVNKPSSERVREIIVDAVKIEQEFTEALPVGLIGMNCILMKQYIEFVADRLVLELGS 300
QY 301 KVFQAE NPFDFMENISLEGKTNFFEKRVSEYQRFVAVMAETTDNVFTLDADF 351
Db 301 KVFQAE NPFDFMENISLEGKTNFFEKRVSEYQRFVAVMAETTDNVFTLDADF 351
RESULT 2
AAB69050
ID AAB69050 standard; protein; 351 AA.
XX AAB69050;
AC AAB69050;
XX 18-APR-2001 (first entry)
DT 18-APR-2001 (first entry)
XX Human ribonucleotide reductase TP53R2H protein sequence SEQ ID NO:1.
DE Human ribonucleotide reductase TP53R2H protein sequence SEQ ID NO:1.
XX

KW Human; ribonucleotide reductase; cancer; DNA repair; p53.
XX Homo sapiens.
XX WO200100799-A1.
XX 04-JAN-2001.
XX 27-JUN-2000; 2000WO-JP004189.
XX 28-JUN-1999; 99JP-00181131.
PR 06-JUL-1999; 99JP-00192391.
PR 21-JAN-2000; 2000JP-00017770.
XX (TAKE) TAKEDA CHEM IND LTD.
PA (NAKA/) NAKAMURA Y.
XX Nakamura Y, Arakawa H, Tanaka H;
PI WPI; 2001-112446/12.
DR N-PSDB; AAE32438.
XX Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
PT diagnosis, treatment and prevention of cancer.
XX Claim 1; Fig 1-3; 102pp; Japanese.
XX The present sequence represents a human ribonucleotide reductase
CC designated TP53R2H. TP53R2H has cytosolic activity. TP53R2H is part of
CC the DNA repair mechanism and its activity is induced by p53. It can be
CC used for the treatment, prevention and diagnosis of a wide range of
CC cancers
XX SQ Sequence 351 AA;
Query Match 100.0%; Score 1821; DB 4; Length 351;
Best Local Similarity 100.0%; Pred. No. 5.2e-179;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGDPERPAAGLDQDERSSTNESEIKSNEEPLLRKSSRRFVFIQYDPDIWKYKQAO 60
Db 1 MGDPERPAAGLDQDERSSTNESEIKSNEEPLLRKSSRRFVFIQYDPDIWKYKQAO 60
QY 61 ASFWTAEVDLSKDLPHWNKLKADSKYFISHILAFPAASDGI VNVNLFVERFSQEVQVPEA 120
Db 61 ASFWTAEVDLSKDLPHWNKLKADSKYFISHILAFPAASDGI VNVNLFVERFSQEVQVPEA 120
QY 121 RCFYGFQI LIENVHSEMYSLIIDTYIRDPKKREFL FNAIETMPYVKKKADWALRIADRK 180
Db 121 RCFYGFQI LIENVHSEMYSLIIDTYIRDPKKREFL FNAIETMPYVKKKADWALRIADRK 180
QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGIMPGLTFSNELISRDEGLHCDPACLMFQ 240
Db 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGIMPGLTFSNELISRDEGLHCDPACLMFQ 240
QY 241 YLVNKPSSERVREIIVDAVKIEQEFTEALPVGLIGMNCILMKQYIEFVADRLVLELGS 300
Db 241 YLVNKPSSERVREIIVDAVKIEQEFTEALPVGLIGMNCILMKQYIEFVADRLVLELGS 300
QY 301 KVFQAE NPFDFMENISLEGKTNFFEKRVSEYQRFVAVMAETTDNVFTLDADF 351
Db 301 KVFQAE NPFDFMENISLEGKTNFFEKRVSEYQRFVAVMAETTDNVFTLDADF 351
RESULT 3
AAB93497
ID AAB93497 standard; protein; 351 AA.
XX AAB93497;
AC AAB93497;
XX 26-JUN-2001 (first entry)
DT 26-JUN-2001 (first entry)
XX Human protein sequence SEQ ID NO:12811.
DE Human protein sequence SEQ ID NO:12811.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX Homo sapiens.
 OS EP1074617-A2.
 PN 07-FEB-2001.
 PD 28-JUL-2000; 2000EP-00116126.
 PF 29-JUL-1999; 99JP-00248036.
 PR 27-AUG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.
 XX (HELI-) HELIX RES INST.
 PA Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WPI; 2001-318749/34.
 DR Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
 XX Claim 8; SEQ ID NO 12811; 2537pp + Sequence Listing; English.
 XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention
 XX Sequence 351 AA;
 SQ
 Query Match 99.8%; Score 1817; DB 4; Length 351;
 Best Local Similarity 99.7%; Pred. NO. 1.4e-178;
 Matches 350; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGDPERPAAGLDDQERSSDTNESEIKSNEBPLLRKSRFRVFPFIQYDPDIWKMYKQAO 60
 DB 1 MGDPERPAAGLDDQERSSDTNESEIKSNEBPLLRKSRFRVFPFIQYDPDIWKMYKQAO 60
 QY 61 ASFWTAEEVDLSKLPWNKLKADKEYFTSHILAPFAASDGVNENLVERFSQVQVPEA 120
 DB 61 ASFWTAEEVDLSKLPWNKLKADKEYFTSHILAPFAASDGVNENLVERFSQVQVPEA 120
 QY 121 RCFYGFQILINNVHSEMSLLIDTYIRDPKREFLFNAITETMPYVKKKADWALRWIADRK 180
 DB 121 RCFYGFQILINNVHSEMSLLIDTYIRDPKREFLFNAITETMPYVKKKADWALRWIADRK 180
 QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKKRGLMPLGTFNSNELISRDEGLHCDFACLMFQ 240

DB 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKKRGLMPLGTFNSNELISRDEGLHCDFACLMFQ 240
 QY 241 YLVNKPSEERREIIVDAVKIEQSEFLTEALPVGLIGNNCILMKQYIEFVADRLVLELGS 300
 DB 241 YLVNKPSEERREIIVDAVKIEQSEFLTEALPVGLIGNNCILMKQYIEFVADRLVLELGS 300
 QY 301 KVFQAEPPDFMENISLEGKTNPFPEKRVSEYQRFVAFVMAETTDNVFTLDADF 351
 DB 301 KVFQAEPPDFMENISLEGKTNPFPEKRVSEYQRFVAFVMAETTDNVFTLDADF 351
 RESULT 4
 AAB69051
 ID AAB69051 standard; protein; 389 AA.
 XX AAB69051;
 AC 18-APR-2001 (first entry)
 XX Human R2 protein sequence SEQ ID NO:5.
 DT Human; ribonucleotide reductase; cancer; DNA repair; p53.
 DE Homo sapiens.
 XX OS
 XX PN WO200100799-A1.
 XX PD 04-JAN-2001.
 XX PF 27-JUN-2000; 2000WO-JP004189.
 XX PR 28-JUN-1999; 99JP-00181131.
 PR 06-JUL-1999; 99JP-00192391.
 PR 21-JAN-2000; 2000JP-00017770.
 XX (TAKE) TAKEDA CHEM IND LTD.
 PA (NAKA/) NAKAMURA Y.
 XX Nakamura Y, Arakawa H, Tanaka H;
 XX WPI; 2001-112446/12.
 DR Ribonucleotide reductase involved in DNA repair and DNA encoding it, for diagnosis, treatment and prevention of cancer.
 PT Example 2; Page 92-94; 102pp; Japanese.
 XX The present invention describes a human ribonucleotide reductase designated TP53R2H. TP53R2H has cytosolic activity. TP53R2H is part of the DNA repair mechanism and its activity is induced by p53. It can be used for the treatment, prevention and diagnosis of a wide range of cancers. The present sequence represents the human R2 protein which is used in an example from the present invention
 CC Sequence 389 AA;
 SQ
 Query Match 79.8%; Score 1454; DB 4; Length 389;
 Best Local Similarity 83.8%; Pred. NO. 4.8e-141;
 Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;
 QY 31 EEPPLRKSRRFVFPFIQYDPDIWKMYKQAOASFWTAEEVDLSKLPWNKLKADKEYIS 90
 DB 69 DEPLLRENPRFRVFPFIETDHIWQYKKAESFWTAEEVDLSKLPWNKLKADKEYIS 128
 QY 91 HILAPFAASDGVNENLVERFSQVQVPEARCFYGFQILINNVHSEMSLLIDTYIRDPK 150
 DB 129 HVLAPFAASDGVNENLVERFSQVQVPEARCFYGFQILINNVHSEMSLLIDTYIRDPK 188
 QY 151 KREFLFNAITETMPYVKKKADWALRWIADRKSTFGERVVAFVMAETTDNVFTLDADF 210
 DB 189 EREFLFNAITETMPYVKKKADWALRWIADRKSTFGERVVAFVMAETTDNVFTLDADF 248

QY 211 RGLMPLTFSNELISRDEGLHCDPACLMFQYLVNKPSEVRREIIVDAVKIQEFLTEAL 270
 DB 249 RGLMPLTFSNELISRDEGLHCDPACLMFQYLVNKPSEVRREIIVDAVKIQEFLTEAL 308
 QY 271 PVGLIGMNCILMKQYIEFVADRLVLLVELGSKVQFAENPDPFNMENISLEGKTNFFEKRVSE 330
 DB 309 PVKLGIMNCTLMKQYIEFVADRLVLLVELGSKVQFAENPDPFNMENISLEGKTNFFEKRVGE 368
 QY 331 YQRFVMAETTDNVFTLDADF 351
 DB 369 YQRMGMSSPTENSFTLDADF 389

RESULT 5
 ID ABU07433 standard; protein; 389 AA.
 AC ABU07433;
 DT 28-JAN-2003 (first entry)
 XX Protein differentially regulated in prostate cancer #36.
 DE Prostate cancer; gene expression; differential regulation;
 KW molecular marker; drug target; cancer detection; cancer diagnosis;
 KW cancer staging; cancer grading; cancer assessing; cancer monitoring.
 XX Homo sapiens.
 OS WO200281638-A2.
 PN 17-OCT-2002.
 PD 08-APR-2002; 2002WO-US010824.
 PF 06-APR-2001; 2001US-0281731P.
 PR 06-APR-2001; 2001US-0281732P.
 XX (ORIG-) ORIGENE TECHNOLOGIES INC.
 PA Sun Z, Jay G;
 PI WPI; 2003-058520/05.
 DR N-PSDB; ABX10335.
 XX Novel genes which are differentially regulated in prostate cancer, useful for diagnosing prostate cancer in prostate tissue sample and assessing therapeutic or preventive intervention in prostate cancer patients.

Claim 1; Page 273-274; 416pp; English.

The invention describes genes (I) which are differentially regulated in prostate cancer. (I) is useful for diagnosing a prostate cancer in a sample comprising prostate tissue, which involves determining the number of target genes which are differentially-regulated in the sample, where the number is indicative of the probability that the sample comprises prostate cancer. (I) is useful for assessing a therapeutic or preventive intervention in a subject having a prostate cancer, which involves determining the expression levels in a sample comprising prostate tissue of target genes which are differentially-regulated in prostate cancer. Preferably, the expression levels of at least 10 genes are determined. (I) is also useful for identifying agents that modulate a biological activity of a polypeptide differentially-regulated in prostate cancer cells, which involves contacting a polypeptide differentially-regulated in prostate cancer cells with a test agent under conditions effective for the test agent to modulate a biological activity of the polypeptide, and determining whether the test agent modulates the biological activity. (I) is useful as molecular markers, as drug targets, and for detecting, diagnosing, staging, grading, assessing, monitoring, prognosticating, preventing or treating, determining predisposition to diseases and conditions especially relating to prostate cancer. (I) and its expression products are used in the diagnostic test to assay for presence of cancer e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in

CC blood etc. (I) is useful for assessing cancer e.g., to determine the type of cancer, its stage of development, the nature of genetic defect, etc.
 CC The polypeptide encoded by (I) can be used as target for therapy or drug discovery. (I) can also be used for expressing the polypeptide and thus for searching specific binding partners of the polypeptide. (I) is useful in therapeutic applications to treat prostate cancer. The identification of specific genes, and groups of genes, expressed in pathways physiologically relevant to prostate cancer permits the definition of functional and disease pathways and the delineation of targets in these pathways which are useful in diagnostic, therapeutic, and clinical applications. This is the amino acid sequence of a protein differentially regulated in prostate cancer
 XX
 SQ Sequence 389 AA;

Query Match 79.8%; Score 1454; DB 6; Length 389;
 Best Local Similarity 83.8%; Pred. No. 4.8e-141;
 Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;

QY 31 EEPLLRKSRFFVIFPIQYDPDIWKMVKQAQASFWTAEEVDLSKDLPHWNKLKADKPYIS 90
 DB 69 DEPLLRNPRFRVIFPIEYHDIMQYKKAESFWTAEEVDLSKDIQHWESLKPERYFIS 128
 QY 91 HILAPFAASDGI VNNELVERPSQEVQVPRARCFYGFQIILNHNSEMYSLIIDYIRDPK 150
 DB 129 HVLAPFAASDGI VNNELVERPSQEVQVPRARCFYGFQIILNHNSEMYSLIIDYIRDPK 188
 QY 151 KREFLNATETMPYKKKADWALRWIADRKSTGERVVAANVEGVFSGSFAAIFWLK 210
 DB 189 EREFLNATETMPYKKKADWALRWIADRKSTGERVVAANVEGVFSGSFAAIFWLK 248
 QY 211 RGLMPLTFSNELISRDEGLHCDPACLMFQYLVNKPSEVRREIIVDAVKIQEFLTEAL 270
 DB 249 RGLMPLTFSNELISRDEGLHCDPACLMFQYLVNKPSEVRREIIVDAVKIQEFLTEAL 308
 QY 271 PVGLIGMNCILMKQYIEFVADRLVLLVELGSKVQFAENPDPFNMENISLEGKTNFFEKRVSE 330
 DB 309 PVKLGIMNCTLMKQYIEFVADRLVLLVELGSKVQFAENPDPFNMENISLEGKTNFFEKRVGE 368
 QY 331 YQRFVMAETTDNVFTLDADF 351
 DB 369 YQRMGMSSPTENSFTLDADF 389

RESULT 6

ADJ68753
 ID ADJ68753 standard; protein; 389 AA.
 AC ADJ68753;
 DT 06-MAY-2004 (first entry)

XX Human heat mitochondrial protein as a therapeutic target SeqID559.
 DE
 XX Mitochondrial; human; screening assay; diabetes mellitus;
 KW Huntington's disease; osteoarthritis; LHON;
 KW Leber's hereditary optic neuropathy;
 KW Mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
 KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
 KW neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;
 KW osteopathic; ophthalmological; cytostatic.
 XX Homo sapiens.
 OS
 XX WO2003087768-A2.
 PN
 XX 23-OCT-2003.
 PD
 XX 04-APR-2003; 2003WO-US010870.
 PF
 XX 12-APR-2002; 2002US-0372843P.
 PR
 XX 17-JUN-2002; 2002US-0389987P.
 PR
 XX 20-SEP-2002; 2002US-0412418P.

XX PA (MITO-) MITOKOR.
PA (BUCK-) BUCK INST AGE RES.
XX PI Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;
PI Warnock DE;
XX XX WPI; 2003-845369/78.
DR
XX
XX Identifying a mitochondrial target for drug screening assays and for
PT treating diseases associated with altered mitochondrial function,
PT comprises detecting a modified polypeptide in a sample and correlating
PT with the disease.
XX PS
XX Claim 1; SEQ ID NO 559; 180pp; English.
XX
XX This invention relates to novel mitochondrial targets that can be used
CC for therapeutic intervention in treating a disease associated with
CC altered mitochondrial function. Specifically, it refers to a method for
CC identifying proteins of the human heart mitochondrial proteome that are
CC useful for drug screening assays, as well as therapeutic targets. The
CC present invention describes a method for identifying such proteins that
CC can be used in the treatment of various diseases associated with altered
CC mitochondrial function including diabetes mellitus, Huntington's disease,
CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
CC compositions have neuroprotective, neurotropic, antidiabetic,
CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and
CC cytoskeletal activities. This polypeptide sequence is a human heart
CC mitochondrial protein of the invention.
XX
XX Sequence 389 AA;
XX
Query Match 79.8%; Score 1454; DB 7; Length 389;
Best Local Similarity 83.8%; Pred. No. 4.8e-141;
Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;
QY 31 EEPLLRKSSRRFVIFPIQYPDIMWKYKQAQASFWTAEEVDLSKDLPHWNKLKADKRYFIS 90
DB 69 DEPLLRENRRFVIFPIEYHDIWQMYKKAASFWTAEEVDLSKDIQHWESLKPPEERYFIS 128
QY 91 HILAFFAASDGI VNNLVERFSQEVQVPEARCFYGFQIILNHNHSEMYSLLDITYIRDPK 150
DB 129 HVLAFPAASDGI VNNLVERFSQEVQVTEARCFYGFQIAMENIHSEMYSLLDITYIKOPK 188
QY 151 KREFLFNAIETMPYVKKKADWALRWIADRKSTFGERVVAFAAVEGVFFSGSFAAIFWLKK 210
DB 189 EREFLFNAIETMPCVKKKADWALRWIGDKATYGERVVAFAAVEGVFFSGSFAIFWLKK 248
QY 211 RGLMPLGLTSPNELISRDEGLHCDPACLMFQYLVNKPSEERVRRIIVDAVKIQEFLTEAL 270
DB 249 RGLMPLGLTSPNELISRDEGLHCDPACLMFQYLVNKPSEERVRRIIVDAVKIQEFLTEAL 308
QY 271 PVGLIGMNCILMKQYIEFVADRLVLLVLFSGSKVQFAENPFDMENISLEKTNFFPKRVGE 330
DB 309 PVKLLIGMNCILMKQYIEFVADRLVLLVLFSGSKVFRVENPFDMENISLEKTNFFPKRVGE 368
QY 331 YQRFVMAETTDNVFTLDADF 351
DB 369 YQRMGMSSPTENSFTLDADF 389
RESULT 7
ID ADJ66564
XX ADJ66564 standard; protein; 389 AA.
XX AC ADJ66564;
XX AC
XX 06-MAY-2004 (first entry)
XX DE Ribonucleoside-diphosphate reductase M2 chain for anti-cancer complex.
XX

KW KW neuroprotective; cytostatic; gene therapy; protein complex;
KW KW cellular network; cancer; neurodegenerative disease; drug target.
OS Homo sapiens.
PN WO2004009622-A2.
XX XX
XX PD 29-JAN-2004.
XX
XX 18-JUL-2003; 2003WO-EP007835.
XX PF
XX 19-JUL-2002; 2002EP-00016109.
PR 19-JUL-2002; 2002EP-00016111.
PR 19-JUL-2002; 2002EP-00016123.
PR 19-JUL-2002; 2002EP-00016128.
PR 22-JUL-2002; 2002EP-00016427.
XX
XX (CELL-) CELLZOME AG.
PA
XX Merino A, Bouwmeester T, Bauer A, Drewes G, Marzioch M, Kruse U;
PI Superti-Furga G, Eberhard D, Ruffner H, Hobson S, Helftenbein G;
PI Cruciat C;
XX
XX WPI; 2004-123372/12.
XX
XX New protein complexes of cellular networks underlying the development of
PT cancer and other diseases, useful for diagnosing and/or treating
PT neurodegenerative diseases or cancer, and in drug screening.
XX
XX Disclosure; SEQ ID NO 94; 809pp; English.
XX
XX The invention relates to a protein complex of cellular networks
CC underlying the development of cancer and other diseases. The complex (I)
CC comprises at least one first and second proteins selected from any of the
CC proteins listed in the specification, or their functionally active
CC derivatives, fragments, homologues or variants, the variants being
CC encoded by a nucleic acid that hybridizes to the nucleic acid encoding
CC the protein under low stringency conditions. A complex (II) comprises at
CC least two of the second proteins, where the low stringency conditions
CC comprise hybridization in a buffer comprising 35% formamide, 5 x SSC, 50
CC mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.02% PVP, 0.02% BSA, 100 microg/ml
CC denatured salmon sperm DNA, and 10% (wt/vol) dextran sulfate for 18-20
CC hours at 40 deg C, washing in a buffer consisting of 2 x SSC, 25 mM Tris-
CC HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS for 1.5 hours at 55 deg C, and
CC washing in a buffer consisting of 2 x SSC, 25 mM Tris-HCl (pH 7.4), 5 mM
CC EDTA, and 0.1% SDS for 1.5 hours at 60 deg C. The composition and methods
CC are useful in diagnosing or treating diseases and disorders, preferably
CC neurodegenerative diseases. These may also be used as a drug target or in
CC manufacturing a medicament for the treatment or prevention of the above-
CC mentioned diseases or disorders. The composition may also be used for
CC treating cancer. This sequence represents one of the proteins of the
CC complex of the invention.
XX
XX Sequence 389 AA;
XX
Query Match 79.8%; Score 1454; DB 8; Length 389;
Best Local Similarity 83.8%; Pred. No. 4.8e-141;
Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;
QY 31 EEPLLRKSSRRFVIFPIQYPDIMWKYKQAQASFWTAEEVDLSKDLPHWNKLKADKRYFIS 90
DB 69 DEPLLRENRRFVIFPIEYHDIWQMYKKAASFWTAEEVDLSKDIQHWESLKPPEERYFIS 128
QY 91 HILAFFAASDGI VNNLVERFSQEVQVPEARCFYGFQIILNHNHSEMYSLLDITYIRDPK 150
DB 129 HVLAFPAASDGI VNNLVERFSQEVQVTEARCFYGFQIAMENIHSEMYSLLDITYIKOPK 188
QY 151 KREFLFNAIETMPYVKKKADWALRWIADRKSTFGERVVAFAAVEGVFFSGSFAAIFWLKK 210
DB 189 EREFLFNAIETMPCVKKKADWALRWIGDKATYGERVVAFAAVEGVFFSGSFAIFWLKK 248
QY 211 RGLMPLGLTSPNELISRDEGLHCDPACLMFQYLVNKPSEERVRRIIVDAVKIQEFLTEAL 270
DB

```

Db      249 RGLMPGLTFSNELISRDEGLHCDPACLMFKHLVHKPSEERVREIIINAVRIEQEFTLEAL 308
QY      271 PVGLIGMNCILMKQYIEFVADRLLLVGLGSKVFOAENPFDPMENISLEGKTNFFEKRVSE 330
Db      309 PVKLLGMNCTLMKQYIEFVADRLMLGLGSKVFRVENPFDPMENISLEGKTNFFEKRVSE 368
QY      331 YQFVMAETTDNVFTLDADF 351
Db      369 YQRMGMSSPTENSFTLDADF 389

RESULT 8
ADK70435
ID      ADK70435 standard; protein; 389 AA.
XX
AC      ADK70435;
XX
DT      06-MAY-2004 (first entry)
XX
DE      Respiratory disease differentially expressed protein #1.
XX
KW      cytostatic; respiratory; antiasthmatic; gene therapy;
KW      differential gene expression; respiratory disorder; lung cancer;
KW      chronic obstructive pulmonary disease; emphysema; asthma.
XX
OS      Homo sapiens.
XX
PN      WO2003101283-A2.
XX
PD      11-DEC-2003.
XX
PF      02-JUN-2003; 2003WO-US017409.
XX
PR      04-JUN-2002; 2002US-0386005P.
XX
PA      (INCY-) INCYTE CORP.
XX
PI      Rickert PK, Kraenow R;
XX
PI      WPI; 2004-042945/04.
XX
DR
XX
PT      New combination comprising cDNAs and proteins that are differentially
PT      expressed in respiratory disorders, useful for diagnosing or treating
PT      respiratory diseases e.g. lung cancer, chronic obstructive pulmonary
PT      diseases or asthma.
XX
PS      Claim 14; SEQ ID NO 171; 343pp; English.
XX
CC      The invention relates to cDNA sequences that are differentially expressed
CC      in respiratory disorders or their complements or encoded proteins. The
CC      cDNAs and proteins are useful for diagnosing, treating or monitoring
CC      treatment of a subject with a respiratory disease including lung cancer,
CC      chronic obstructive pulmonary diseases, emphysema or asthma. The protein
CC      is also useful for screening molecules or compounds to identify at least
CC      one ligand which specifically binds the protein. It is also useful for
CC      preparing and purifying a polyclonal or monoclonal antibody. This
CC      sequence corresponds to a protein of the invention.
XX
SQ      Sequence 389 AA;

Query Match          79.8%; Score 1454; DB 8; Length 389;
Best Local Similarity 83.8%; Pred. No. 4.8e-141;
Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;

QY      31 EEPLLKSSRRVFIPFIQYPDIMWKYKQAQASFWTAEEVDLSKDLPHWNKLKADEKYFIS 90
Db      69 DEPLLRENPRFVIFPIEYHDIMQYKKAESFWTAEEVDLSKDIQHWESLKPEERYFIS 128
QY      91 HILAFPAASDGIWNLVRFQEVQVPEARCFYGFQIILNVHSEMYSLLLIDTYIRDPK 150
Db      129 HVLAFPAASDGIWNLVRFQEVQVTEARCFYGFQIAMENIHSEMYSLLLIDTYIKDPK 188
QY      151 KREFLFNAIETMPYVKKADWALRWIADRKSTFGERVVFAAVEGVFSGSFAAIFWLKK 210

Db      189 EREFLNAIETMPYVKKADWALRWIGKEATYGERVVFAAVEGVFSGSFAAIFWLKK 248
QY      211 RGLMPGLTFSNELISRDEGLHCDPACLMFOYLVNKPSEERVREIIIVDAVKTEQEFLEAL 270
Db      249 RGLMPGLTFSNELISRDEGLHCDPACLMFKHLVHKPSEERVREIIINAVRIEQEFTLEAL 308
QY      271 PVGLIGMNCILMKQYIEFVADRLLLVGLGSKVFOAENPFDPMENISLEGKTNFFEKRVSE 330
Db      309 PVKLLGMNCTLMKQYIEFVADRLMLGLGSKVFRVENPFDPMENISLEGKTNFFEKRVSE 368
QY      331 YQFVMAETTDNVFTLDADF 351
Db      369 YQRMGMSSPTENSFTLDADF 389

RESULT 9
ADN03789
ID      ADN03789 standard; protein; 389 AA.
XX
AC      ADN03789;
XX
DT      01-JUL-2004 (first entry)
XX
DE      Antipsoriatic protein sequence #91.
XX
KW      antipsoriatic; gene therapy; psoriasis; diagnosis.
XX
OS      Homo sapiens.
XX
PN      WO2004028479-A2.
XX
PD      08-APR-2004.
XX
PF      25-SEP-2003; 2003WO-US030907.
XX
PR      25-SEP-2002; 2002US-0414006P.
XX
XX
PA      (GETH ) GENENTECH INC.
XX
PI      Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
PI      Wu TD;
XX
XX      WPI; 2004-305105/28.
DR      N-PSDB; ADN03788.
XX
XX      New PRO nucleic acid or polypeptide, useful for preparing a
PT      pharmaceutical composition for diagnosing or treating psoriasis in a
PT      mammal.
XX
PS      Claim 9; SEQ ID NO 183; 3069pp; English.
XX
CC      The invention relates to novel polynucleotide and polypeptides for
CC      treating psoriasis or a sequence having at least 80% identity to the
CC      above sequences. The nucleic acid is useful for preparing a composition
CC      for diagnosing or treating psoriasis in a mammal. This sequence
CC      corresponds to one of the polypeptides of the invention.
XX
SQ      Sequence 389 AA;

Query Match          79.8%; Score 1454; DB 8; Length 389;
Best Local Similarity 83.8%; Pred. No. 4.8e-141;
Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;

QY      31 EEPLLKSSRRVFIPFIQYPDIMWKYKQAQASFWTAEEVDLSKDLPHWNKLKADEKYFIS 90
Db      69 DEPLLRENPRFVIFPIEYHDIMQYKKAESFWTAEEVDLSKDIQHWESLKPEERYFIS 128
QY      91 HILAFPAASDGIWNLVRFQEVQVPEARCFYGFQIILNVHSEMYSLLLIDTYIRDPK 150
Db      129 HVLAFPAASDGIWNLVRFQEVQVTEARCFYGFQIAMENIHSEMYSLLLIDTYIKDPK 188
QY      151 KREFLFNAIETMPYVKKADWALRWIADRKSTFGERVVFAAVEGVFSGSFAAIFWLKK 210

```

Db 189 EREFLNAIETMPCVKKKADWALRWIGDKYATYGERVVAFAAVEGIFSGSFASIFWLKK 248
Qy 211 RGLMPGLTFSNELISRDEGLHCDFACLMFOYLNVKPSSEERVRRIIIVDAVKIQEFLTEAL 270
Db 249 RGLMPGLTFSNELISRDEGLHCDFACLMFKHLVHKPSSEERVRRIIINAVRIQEFLTEAL 308
Qy 271 PVGLIGNMCTLMKQYIEFVADRLVDELGFSKVFOAENPFDPMENISLEGKTNFFEKRVSE 330
Db 309 PVKLGNNCTLMKQYIEFVADRLMELGFSKVRVENPFDPMENISLEGKTNFFEKRVGE 368
Qy 331 YORFVMAETTDNVFTLDADF 351
Db 369 YQRMGMVMSPTENSFTLDADF 389
RESULT 10
ADN04444
ID ADN04444 standard; protein; 389 AA.
XX AC ADN04444;
XX DT 01-JUL-2004 (first entry)
XX DE Antipsoriatic protein sequence #415.
XX KW antipsoriatic; gene therapy; psoriasis; diagnosis.
XX OS Homo sapiens.
XX PN WO2004028479-A2.
XX PD 08-APR-2004.
XX PF 25-SEP-2003; 2003WO-US030907.
XX PR 25-SEP-2002; 2002US-0414006P.
XX PA (GETH) GENENTECH INC.
XX PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
PI Wu TD;
XX WPI: 2004-305105/28.
XX DR N-PSDB; ADN04443.
XX PT New PRO nucleic acid or polypeptide, useful for preparing a
PT pharmaceutical composition for diagnosing or treating psoriasis in a
XX mammal.
XX PS Claim 9; SEQ ID NO 838; 3069pp; English.
XX CC The invention relates to novel polynucleotide and polypeptides for
CC treating psoriasis or a sequence having at least 80% identity to the
CC above sequences. The nucleic acid is useful for preparing a composition
CC for diagnosing or treating psoriasis in a mammal. This sequence
XX corresponds to one of the polypeptides of the invention.
XX SQ Sequence 389 AA;
Query Match 79.8%; Score 1454; DB 8; Length 389;
Best Local Similarity 83.8%; Pred. No. 4.8e-141;
Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;
Qy 31 BEPLLRKSRFVPIPIQVPOIKWYKQAQSFWTAEVDLSKDLPHWNKLKADKEYFIS 90
Db 69 DEPLLRNPRFVPIPIEYHDIWQYKKAESFWTAEVDLSKDLQHWESLPREYFIS 128
Qy 91 HILAFFAASDGVNENLVERFSQEQVPEARCFCYQFQILINVHSEMYSLLDITVIRDPK 150
Db 129 HVLAFFAASDGVNENLVERFSQEQVITEARCFCYQFQIAMENIHSEMYSLLDITVIKDPK 188
Qy 151 KREFLFNAIETWPKYKADWALRWIADRKSTFGSERVVAFAAVEGIFSGSFASIFWLKK 210

Db 189 EREFLNAIETMPCVKKKADWALRWIGDKYATYGERVVAFAAVEGIFSGSFASIFWLKK 248
Qy 211 RGLMPGLTFSNELISRDEGLHCDFACLMFOYLNVKPSSEERVRRIIIVDAVKIQEFLTEAL 270
Db 249 RGLMPGLTFSNELISRDEGLHCDFACLMFKHLVHKPSSEERVRRIIINAVRIQEFLTEAL 308
Qy 271 PVGLIGNMCTLMKQYIEFVADRLVDELGFSKVFOAENPFDPMENISLEGKTNFFEKRVSE 330
Db 309 PVKLGNNCTLMKQYIEFVADRLMELGFSKVRVENPFDPMENISLEGKTNFFEKRVGE 368
Qy 331 YORFVMAETTDNVFTLDADF 351
Db 369 YQRMGMVMSPTENSFTLDADF 389
RESULT 11
ADO19226
ID ADO19226 standard; protein; 389 AA.
XX AC ADO19226;
XX DT 12-AUG-2004 (first entry)
XX DE Human PRO polypeptide #79.
XX KW Human; PRO; immune related disorder; systemic lupus erythematosus;
KW rheumatoid arthritis; osteoarthritis; juvenile chronic arthritis;
KW systemic sclerosis; Sjogren's syndrome; vasculitis; sarcoidosis;
KW autoimmune haemolytic anaemia; autoimmune thrombocytopenia; thyroiditis;
KW diabetes mellitus; renal disease; demyelinating disease;
KW central nervous system; peripheral nervous system;
KW demyelinating polyneuropathy; Guillain-Barre syndrome;
KW chronic inflammatory demyelinating polyneuropathy.
XX OS Homo sapiens.
XX PN WO2004043361-A2.
XX PD 27-MAY-2004.
XX PF 06-NOV-2003; 2003WO-US035268.
XX PR 08-NOV-2002; 2002US-0425235P.
XX PA (GETH) GENENTECH INC.
XX PI Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM;
PI Wood WI, Wu TD;
XX WPI: 2004-420067/39.
XX DR N-PSDB; ADO19225.
XX CC Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for
XX treating an immune related disorder such as systemic lupus erythematosus,
XX rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or
XX spondyloarthropathy.
XX PS Claim 7; SEQ ID NO 158; 1731pp; English.
XX CC The invention relates to human PRO polypeptides and the polynucleotides
XX encoding them. The polypeptides and polynucleotides are useful for
XX treating and diagnosing immune related disorders in mammals. The immune
XX related disorders include systemic lupus erythematosus, rheumatoid
XX arthritis, osteoarthritis, juvenile chronic arthritis, systemic
XX sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune
XX haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes
XX mellitus, immune-mediated renal disease, demyelinating diseases of the
XX central or peripheral nervous system, demyelinating polyneuropathy,
XX Guillain-Barre syndrome and chronic inflammatory demyelinating
XX polyneuropathy. This sequence represents a human PRO polypeptide of the
XX invention.


```
SQ Sequence 389 AA;
Query Match 79.8%; Score 1454; DB 8; Length 389;
Best Local Similarity 83.8%; Pred. No. 4.8e-141;
Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;
QY 31 EEP LKSR RRV FIPPIQY PDIW MYKQAQASF WTAEEVDLSKDLPHWNK LKAD EKYFIS 90
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
69 DE P L L R E N P R R F V I P I E Y H D I W Q Y K K A E A S F W T A E E V D L S K D I Q H W E S L K P E E R Y F I S 128
QY 91 H I L A F P A A S D G I V N E L V E R F S Q E V Q V P E A R C F Y G F Q I L I E N V H S E M Y S L L I D T Y I R D P K 150
Db H V L A F P A A S D G I V N E L V E R F S Q E V Q V P E A R C F Y G F Q I L I E N V H S E M Y S L L I D T Y I R D P K 188
QY 151 K R E F L F N A I E T M P V Y K K A D W A L R W I A D R K S T F G R V V A F A A V E G V F F S G S F A A I F W L K K 210
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
189 E R E F L F N A I E T M P C V K K A D W A L R W I G D K E A T Y G E R V V A F A A V E G I F F S G S F A S I F W L K K 248
QY 211 R G L M P G L T F S N E L I S R D E G L H C D F A C L M F Q Y L V N K P S E R V R E I I V D A V K I Q E F L T E A L 270
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
249 R G L M P G L T F S N E L I S R D E G L H C D F A C L M F K H L V H K P S E R V R E I I N A V R I E Q E F L T E A L 308
QY 271 P V G L I G M N C T L M K Q I E F V A D R L L V E L G F S K V F Q A E N P D F M E N I S L G K T N F F E K R V G E 330
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
309 P V K L I G M N C T L M K Q I E F V A D R L L V E L G F S K V F R V E N P D F M E N I S L G K T N F F E K R V G E 368
QY 331 Y Q R F A V M A E T T D N V T L D A D F 351
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
369 Y Q R M G M S P T E N S F T L D A D F 389

RESULT 12
ADQ09272
ID ADQ09272 standard; protein; 389 AA.
XX
AC ADQ09272;
XX
DT 23-SEP-2004 (first entry)
XX
DE Human RRM2 protein SEQ ID NO:457.
XX
KW thanatos-associated protein; THAP; THAP responsive gene; THAP family;
KW THAP responsive element; angiogenesis; inflammation; metastasis; cancer;
KW apoptosis; cardiovascular disease; neurodegenerative disease; chemokine;
KW antiangiogenic; antiinflammatory; cardiovascular; cytostatic;
KW neuroprotective; osteopathic; THAP modulator; THAP synthesis modulator;
KW human.
XX
OS Homo sapiens.
XX
WO2004055050-A2.
XX
PD 01-JUL-2004.
XX
PF 10-DEC-2003; 2003WO-IB006434.
XX
PR 10-DEC-2002; 2002US-0432699P.
XX
PR 03-JUL-2003; 2003US-0485027P.
XX
PA (ENDO-) ENDOCUBE SAS.
XX
PA (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI Girard J, Amalric F, Roussigne M, Clouaire T;
XX
DR WPI; 2004-525034/50.
XX
DR N-PSDB; ADQ09273.
XX
PT Modulating expression of a Thanatos (death)-Associated Protein (THAP)
PT responsive gene for preventing or treating e.g. cancer or inflammation,
PT comprises modulating the interaction of a THAP polypeptide with a nucleic
PT acid.
XX
PS Example 47; SEQ ID NO 457; 612pp; English.
```

```
XX
CC The present invention describes a method for modulating the expression of
CC a thanatos (death)-associated protein (THAP) responsive gene. The method
CC comprises modulating the interaction of a THAP-family polypeptide or its
CC biological fragment with a nucleic acid, and so enhancing or repressing
CC the expression of the THAP responsive gene. Also described: (1) a method
CC of modulating the expression of a gene responsive to a THAP/chemokine
CC complex; (2) a pharmaceutical composition comprising a THAP responsive
CC element in a pharmaceutical carrier; (3) a transcription factor decoy
CC consisting essentially of a THAP responsive element; (4) a cell
CC comprising a transcription factor decoy described above; (5) methods of
CC modulating the interaction between a nucleic acid and a THAP-family
CC polypeptide or its biological fragment, or a nucleic acid and a
CC THAP/chemokine complex; (6) a vector packaging cell line comprising a
CC cell comprising a viral vector which comprises a promoter operably linked
CC to a nucleic acid encoding a THAP-family polypeptide or its biological
CC fragment; (7) a method of constructing a cell which expresses a
CC recombinant THAP-family polypeptide; (8) a method of ameliorating
CC symptoms associated with a condition mediated by a THAP/chemokine complex
CC; (9) methods of identifying a test compound that modulates transcription
CC at a THAP responsive element or that modulates the transport of a
CC chemokine into the nucleus; (10) methods for reducing the symptoms
CC associated with a condition selected from excessive or insufficient
CC angiogenesis, inflammation, metastasis of a cancerous tissue, excessive
CC or insufficient apoptosis, cardiovascular disease and neurodegenerative
CC diseases; symptoms associated with a condition resulting from the
CC activity of a chemokine or a THAP-family polypeptide in an individual; or
CC symptoms associated with transcriptional repression or activation
CC mediated by a THAP-family polypeptide in an individual; (11) a vector
CC encoding a THAP responsive promoter operably linked to a nucleic acid
CC encoding a detectable product; (12) a genetically engineered cell
CC comprising the vector described above or that expresses a THAP-family
CC polypeptide or its biological fragment; (13) an in vitro transcription
CC reaction comprising a nucleic acid comprising a THAP responsive promoter,
CC ribonucleotides and an RNA polymerase; and (14) an isolated mutant THAP-
CC family polypeptide that does not bind to a chemokine. The pharmaceutical
CC composition has antiangiogenic, antiinflammatory, cardiovascular,
CC cytostatic, neuroprotective and osteopathic activities, and can be used
CC as a THAP and THAP synthesis modulator. The composition can be used for
CC modulating the expression of a THAP responsive gene. Modulation is useful
CC for reducing symptoms of conditions such as excessive or insufficient
CC angiogenesis, inflammation, metastasis of a cancerous tissue, excessive
CC or insufficient apoptosis, cardiovascular disease or neurodegenerative
CC diseases. The present sequence is used in the exemplification of the
CC present invention.
XX
SQ Sequence 389 AA;
```

```
Query Match 79.8%; Score 1454; DB 8; Length 389;
Best Local Similarity 83.8%; Pred. No. 4.8e-141;
Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;
```

```
QY 31 EEP LKSR RRV FIPPIQY PDIW MYKQAQASF WTAEEVDLSKDLPHWNK LKAD EKYFIS 90
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
69 DE P L L R E N P R R F V I P I E Y H D I W Q Y K K A E A S F W T A E E V D L S K D I Q H W E S L K P E E R Y F I S 128
QY 91 H I L A F P A A S D G I V N E L V E R F S Q E V Q V P E A R C F Y G F Q I L I E N V H S E M Y S L L I D T Y I R D P K 150
Db H V L A F P A A S D G I V N E L V E R F S Q E V Q V P E A R C F Y G F Q I L I E N V H S E M Y S L L I D T Y I R D P K 188
QY 151 K R E F L F N A I E T M P V Y K K A D W A L R W I A D R K S T F G R V V A F A A V E G V F F S G S F A A I F W L K K 210
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
189 E R E F L F N A I E T M P C V K K A D W A L R W I G D K E A T Y G E R V V A F A A V E G I F F S G S F A S I F W L K K 248
QY 211 R G L M P G L T F S N E L I S R D E G L H C D F A C L M F Q Y L V N K P S E R V R E I I V D A V K I Q E F L T E A L 270
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
249 R G L M P G L T F S N E L I S R D E G L H C D F A C L M F K H L V H K P S E R V R E I I N A V R I E Q E F L T E A L 308
QY 271 P V G L I G M N C T L M K Q I E F V A D R L L V E L G F S K V F Q A E N P D F M E N I S L G K T N F F E K R V G E 330
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
309 P V K L I G M N C T L M K Q I E F V A D R L L V E L G F S K V F R V E N P D F M E N I S L G K T N F F E K R V G E 368
QY 331 Y Q R F A V M A E T T D N V T L D A D F 351
```

Query Match 79.8%; Score 1454; DB 8; Length 389;
Best Local Similarity 83.8%; Pred. No. 4.8e-141;
Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;

Db 373 PVKLGMMCTLMKQYIEFVADRLMLELGFSGKVRVENPFDPMENISLEGKTNFFEKRVGE 432
Qy 331 YORFAVMAETTDNVFTLDADF 351
Db 433 YORMGVMSPTENSFTLDADF 453

Search completed: October 30, 2005, 06:56:17
Job time : 75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2005, 06:48:11 ; Search time 21 seconds
(without alignments)
1608.194 Million cell updates/sec

Title: US-10-698-228-1
Perfect score: 1821
Sequence: 1 MGDPERPEAGLDQDERSSS.....QRFVAVAEITDNTVFTLDADF 351

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1821	100.0	366	2 T46249	hypothetical prote
2	1454	79.8	389	2 S25854	ribonucleoside-dip
3	1443	79.2	390	1 S05735	ribonucleoside-dip
4	1421	78.0	386	2 S27153	ribonucleoside-dip
5	1331	73.1	384	2 S24585	ribonucleoside-dip
6	1280	70.3	319	2 T30782	ribonucleoside-dip
7	1280	70.3	319	2 B72154	B4L protein - vari
8	1279	70.2	319	1 RDVZVV	ribonucleoside-dip
9	1279	70.2	333	2 T28466	ribonucleoside-dip
10	1266	69.5	333	2 H36839	C8L protein - vari
11	1258.5	69.1	348	2 T30470	ribonucleoside-dip
12	1217	66.8	299	1 RDSS2R	ribonucleoside-dip
13	1214	66.7	381	2 T19876	hypothetical prote
14	1168.5	64.2	399	1 A26916	ribonucleoside-dip
15	1161.5	63.8	391	2 T39992	ribonucleoside-dip
16	1161.5	63.8	391	2 S34808	ribonucleoside-dip
17	1128.5	62.0	329	2 T03688	ribonucleoside-dip
18	1070.5	58.8	340	2 S68538	ribonucleoside-dip
19	1024.5	56.3	349	2 B49412	ribonucleoside-dip
20	950	52.2	324	2 T17978	probable ribonucle
21	805	44.2	345	2 S59744	ribonucleoside-dip
22	681	37.4	668	2 T29884	hypothetical prote
23	638	35.0	327	1 RDVZAS	ribonucleoside-dip
24	385	21.1	331	2 A84389	ribonucleoside red
25	384.5	21.1	415	2 A83502	ribonucleoside red
26	358.5	19.7	346	2 G81728	ribonucleoside-dip
27	357.5	19.6	346	2 E71466	probable ribonucle
28	335	18.4	346	2 B72010	ribonucleoside-dip
29	335	18.4	346	2 F86613	ribonucleoside red

ALIGNMENTS

RESULT 1

T46249

hypothetical protein DKFzp761E1312.1 - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004

C;Accession: T46249

R;Ansorge, W.; Wirkner, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, January 2000

A;Reference number: Z23028

A;Accession: T46249

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-366 <AAA>

A;Cross-references: UNIPROT:Q9NTD8; EMBL:AL137348

A;Experimental source: adult amygdala; clone DKFzp761E1312

C;Genetics:

A;Note: DKFzp761E1312.1

C;Superfamily: ribonucleoside reductase small subunit

Query Match 100.0%; Score 1821; DB 2; Length 366;
Best Local Similarity 100.0%; Pred. No. 1.6e-141;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MGDPPEAAGLDQDERSSSDTNESEIKSNEEPLLRKSSRRFVFPPIQYDPIWKMYKQAAQ	60
Db	16	MGDPPEAAGLDQDERSSSDTNESEIKSNEEPLLRKSSRRFVFPPIQYDPIWKMYKQAAQ	75
Qy	61	ASFWTAEVDLSKDLPHWNKLGKADKDYFISHILAFFAASDGIIVNENLVERFSQEVQVPEA	120
Db	76	ASFWTAEVDLSKDLPHWNKLGKADKDYFISHILAFFAASDGIIVNENLVERFSQEVQVPEA	135
Qy	121	RCFYGQILTIENVHSEMYSLIDTYIRDPKKRELFNAIETMPYVKKADWALRIADRK	180
Db	136	RCFYGQILTIENVHSEMYSLIDTYIRDPKKRELFNAIETMPYVKKADWALRIADRK	195
Qy	181	STFGERVAFAAVEGVFFSGSFAAIFWLKKRGLMPGLTFSNELISRDEGLHCDFACLMFQ	240
Db	196	STFGERVAFAAVEGVFFSGSFAAIFWLKKRGLMPGLTFSNELISRDEGLHCDFACLMFQ	255
Qy	241	YLKNSSEERVRRIIVDAVKIEQFLTEALPVGLIGNLCILMKQYIEFVADRLVGLGFS	300
Db	256	YLKNSSEERVRRIIVDAVKIEQFLTEALPVGLIGNLCILMKQYIEFVADRLVGLGFS	315
Qy	301	KVFAENPFDFMENISLEGKTNFPEKRVSEYQRFVAVMAETDNTVFTLDADF	351
Db	316	KVFAENPFDFMENISLEGKTNFPEKRVSEYQRFVAVMAETDNTVFTLDADF	366

RESULT 2

S25854

ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain - human

N;Alternate names: ribonucleotide reductase M2 chain; ribonucleotide reductase small cha


```
QY 200 GSPAAITFWLKKRGLMPLGTFNSNELISRDEGLHCDFAFLMPQYLVNKPSEBVRREIIVDAV 259
Db 239 GSPASIFWLKKRGLMPLGTFNSNELISRDEGLHCDFAFLMPKHLVHKPSEBVRREIIVDAV 298
QY 260 KIEQELTEALPVGLIGMNCILMKQYIEFVADRLVLVGLGFSKVFQAEVDFPDMENISLEG 319
Db 299 RIEQELTEALPVGLIGMNCILMKQYIEFVADRLVLVGLGFSKVFQAEVDFPDMENISLEG 358
QY 320 KTNFFPKRYSEYQRFVMAETDNNVTLDADF 351
Db 359 KTNFFPKRYSEYQRFVMAETDNNVTLDADF 386

RESULT 5
S24585
ribonucleoside-diphosphate reductase (EC 1.17.4.1) chain M2 - Atlantic surf clam
N;Alternate names: ibonucleoside-diphosphate reductase small chain
C;Species: Spisula solidissima (Atlantic surf clam)
C;Date: 22-Nov-1999 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: S24585
R;Hunt, T.; Standart, N.
submitted to the EMBL Data Library, November 1990
A;Reference number: S24584
A;Accession: S24585
A;Molecule type: mRNA
A;Residues: 1-384 <HUN>
A;Cross-references: UNIPROT:P07201; EMBL:X55125; NID:g10341; PIDN:CAA38919.1; PID:g10342
C;Superfamily: ribonucleoside reductase small subunit
C;Keywords: oxidoreductase

Query Match 73.1%; Score 1331; DB 2; Length 384;
Best Local Similarity 74.3%; Pred. No. 2.4e-101; Mismatches 41; Indels 10; Gaps 4;
Matches 257; Conservative 38;

QY 16 ERSSTDN-ESEIK-----SNEEPLRKSRFRVIFPIQYDINWKYQAQAFWTAEE 68
Db 39 QRSTQKTLKQELKPVVKSQQVEPLADNPRFVLPVLPQYHDINWKYKKAASFWTAEE 98
QY 69 VLSKDLPHNKLKADKQYFISHILAFPAASDGINVENLVERPSQVQVPEARCFYGFQI 128
Db 99 VLSKDMAHWSLKKERHFISHVLAFPAASDGINVENLVERPSQVQVTEARCFYGFQI 158
QY 129 LIENHSEMYSLIDTVIRDPKREPLENAETMPYVKKKADWALRWIADRKSTFGERVV 188
Db 159 AMENIHSEMYSLIDTVIRDPKREPLENAETMPYVKKKADWALRWIADRKSTFGERVV 218
QY 189 AFAAEGVFFSGSFAAIFWLKKRGLMPLGTFNSNELISRDEGLHCDFAFLMPQYLVNKPSE 248
Db 219 AFAAEGVFFSGSFAAIFWLKKRGLMPLGTFNSNELISRDEGLHCDFAFLMPQYLVNKPSE 278
QY 249 ERVREIIVDAVKIEQELTEALPVGLIGMNCILMKQYIEFVADRLVLVGLGFSKVFQAEV 308
Db 279 ERHIIQIDEAVKIEQELTEALPVGLIGMNCILMKQYIEFVADRLVLVGLGFSKVFQAEV 338
QY 309 FPFMENISLEGKTNFFPKRYSEYQRFVMAETDNNVTLDADF 351
Db 339 FPFMEHISLEGKTNFFPKRYSEYQRFVMAETDNNVTLDADF 384

RESULT 6
T30782
ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain - vaccinia virus (strain
N;Alternate names: ribonucleotide reductase, small subunit
C;Species: vaccinia virus
A;Variety: strain Ankara
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T30782
R;Antoine, G.; Scheiflinger, F.; Falkner, F.G.; Dorner, F.
submitted to the EMBL Data Library, March 1997
A;Description: The complete genomic sequence of the Modified Vaccinia Ankara (MVA) strain
A;Reference number: Z20877
A;Accession: T30782
A;Status: preliminary; translated from GB/EMBL/DBJ
```

```
A;Molecule type: DNA
A;Residues: 1-319 <ANT>
A;Cross-references: UNIPROT:O57175; EMBL:U94848; PIDN:AAB96415.1
A;Experimental source: strain Ankara
C;Genetics:
A;Note: MVA032L
C;Function:
A;Description: catalyzes the reduction of ribonucleoside diphosphate to deoxyribonucleoside
C;Superfamily: ribonucleoside reductase small subunit
C;Keywords: deoxyribonucleotide biosynthesis; oxidoreductase

Query Match 70.3%; Score 1280; DB 2; Length 319;
Best Local Similarity 74.7%; Pred. No. 2.8e-97;
Matches 239; Conservative 34; Mismatches 45; Indels 2; Gaps 2;

QY 32 EPLLKSRFRVIFPIQYDINWKYQAQAFWTAEEVDLSKDLPHNKLKADKQYFISH 91
Db 2 EPILAKNPRFVIFPIQYHDINWKYKKAASFWTVEVDISKDINDNKLTPDEKYFIKH 61
QY 92 ILAFPAASDGINVENLVERPSQVQVPEARCFYGFQILNENHSEMYSLIDTVIRDPKK 151
Db 62 VLAFPAASDGINVENLAERFCTEQTEARCFYGFQMAENIHSEMYSLIDTVIRDPKK 121
QY 152 REPLFNAETMPYVKKKADWALRWIADRKSTFGERVVAFVAAVEGVFFSGSFAAIFWLKKR 211
Db 122 KNYLFNAETMPYVKKKADWALRWIADRKSTFGERVVAFVAAVEGVFFSGSFAAIFWLKKR 180
QY 212 GLMPGLTFSNELISRDEGLHCDFAFLMPQYLVNKPSEBVRREIIVDAVKIEQELTEALP 271
Db 181 GLMPGLTFSNELISRDEGLHCDFAFLMPQYLVNKPSEBVRREIIVDAVKIEQELTEALP 240
QY 272 VGLIGMNCILMKQYIEFVADRLVLVGLGFSKVFQAEVDFPDMENISLEGKTNFFPKRYSEY 331
Db 241 VKLIGMNCILMKQYIEFVADRLVLVGLGFSKVFQAEVDFPDMENISLEGKTNFFPKRYSEY 300
QY 332 QRPAVMAETDNNVTLDADF 351
Db 301 QRKGVMSQ-EDNHFSLDVDF 319

RESULT 7
B72154
E4L protein - variola minor virus (strain Garcia-1966)
C;Species: variola minor virus
C;Date: 24-Nov-1999 #sequence_revision 24-Nov-1999 #text_change 09-Jul-2004
C;Accession: B72154
R;Shchelkunov, S.N.; Totmenin, A.V.; Gutorov, V.V.; Saifonov, P.F.; Massung, R.F.; Loparev,
submitted to GenBank, March 1998
A;Description: Analysis of the complete coding sequence of DNA of alastrim variola minor
A;Reference number: A72150
A;Accession: B72154
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-319 <SHC>
A;Cross-references: UNIPROT:Q89087; GB:Y16780; NID:g5830555; PIDN:CA854628.1; PID:g5830555
A;Experimental source: strain Garcia-1966
C;Genetics:
A;Gene: E4L
C;Superfamily: ribonucleoside reductase small subunit

Query Match 70.3%; Score 1280; DB 2; Length 319;
Best Local Similarity 75.0%; Pred. No. 2.8e-97;
Matches 240; Conservative 33; Mismatches 45; Indels 2; Gaps 2;

QY 32 EPLLKSRFRVIFPIQYDINWKYQAQAFWTAEEVDLSKDLPHNKLKADKQYFISH 91
Db 2 EPILAKNPRFVIFPIQYHDINWKYKKAASFWTVEVDISKDINDNKLTPDEKYFIKH 61
QY 92 ILAFPAASDGINVENLVERPSQVQVPEARCFYGFQILNENHSEMYSLIDTVIRDPKK 151
Db 62 VLAFPAASDGINVENLAERFCTEQTEARCFYGFQMAENIHSEMYSLIDTVIRDPKK 121
QY 152 REPLFNAETMPYVKKKADWALRWIADRKSTFGERVVAFVAAVEGVFFSGSFAAIFWLKKR 211
```

Db 122 KNYLFNAIETMPCKVKKADWAQKWIHD-SAGYGERLIAFAAAGVIFGSGSFASIFWLKXR 180
QY 212 GLMPGLTFSNELSRDEGLHCDFACLMFQYLVNKPSEERVREIIVDAVKIEQEFLEALP 271
Db 181 GLMPGLTFSNELSRDEGLHCDFACLMFQYLVNKPSEERVRSIITDAVSIEQEFLEALP 240
QY 272 VGLIGNCILMKQYIEFVADRLVELGFSKVFQAEENPDMENISLEGKTNFFEKRVSEY 331
Db 241 VKLIGNCENMKTYIEFVADRLVELGFSKVFQAEENPDMENISLEGKTNFFEKRVSEY 300
QY 332 QRFVMAETTDNVFTLDADF 351
Db 301 QKGMVMSQ-EDNHFSLDVDF 319
RESULT 8
RDVZVW
ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain - vaccinia virus
N;Alternate names: F4L protein
C;Species: vaccinia virus
C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 09-Jul-2004
C;Accession: A29892; I42506; F36213
R;Slabaugh, M.; Roseman, N.; Davis, R.; Mathews, C.
J. Virol. 62, 519-527, 1988
A;Title: Vaccinia virus-encoded ribonucleotide reductase: sequence conservation of the
A;Reference number: A29892; MUID:88091062; PMID:2826813
A;Accession: A29892
A;Molecule type: DNA
A;Residues: 1-319 <SLA>
A;Cross-references: UNIPROT:P11159; GB:M19117; NID:G335808; PID:AAA88680.1; PID:G335809
A;Experimental source: strain WR
R;Goebel, S.J.; Johnson, G.P.; Perkus, M.E.; Davis, S.W.; Winslow, J.P.; Paolletti, E.
Virology 179, 517-563, 1990
A;Title: Appendix to "The complete DNA sequence of vaccinia virus".
A;Reference number: A42501
A;Accession: I42506
A;Molecule type: DNA
A;Residues: 1-212, 'Y', 214-319 <GOB>
A;Cross-references: GB:M35027; NID:G335317; PID:AAA48018.1; PID:G335366
A;Experimental source: strain Copenhagen
R;Goebel, S.J.; Johnson, G.P.; Perkus, M.E.; Davis, S.W.; Winslow, J.P.; Paolletti, E.
Virology 179, 247-266, 1990
A;Title: The complete DNA sequence of vaccinia virus.
A;Reference number: A42531; MUID:91021027; PMID:2219722
A;Contents: annotation; possible protein-coding frames
A;Note: neither amino acid nor nucleotide sequence is given
R;Roseman, N.A.; Slabaugh, M.B.
Virology 178, 410-418, 1990
A;Title: The vaccinia virus HindIII F fragment: nucleotide sequence of the left 6.2 kb.
A;Reference number: A36213; MUID:91020979; PMID:2219701
A;Accession: F36213
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-319 <ROS>
A;Cross-references: EMBL:M34358; NID:G335618; PID:AAA48244.1; PID:G335624
A;Experimental source: strain WR
C;Function:
A;Description: catalyzes the reduction of ribonucleoside diphosphate to deoxyribonucleoside
C;Superfamily: ribonucleoside reductase small subunit
C;Keywords: DNA replication; iron; metalloprotein; oxidoreductase
F;70,101,104,163,197,200/Binding site: 2Fe-O cluster (Asp, Glu, His, Glu, His) #sta
F;108/Active site: Tyr (stable tyrosyl radical) #status predicted
Query Match 70.2%; Score 1279; DB 1; Length 319;
Best Local Similarity 74.7%; Pred. No. 3.3e-97;
Matches 239; Conservative 34; Mismatches 45; Indels 2; Gaps 2;
QY 32 EPLLKSSRRFVFPPIQYDINWYKQAQASFTWAEVDLSKDLPHNKLKADKYPFISH 91
Db 2 EPLIAPNPNRFVFPPIQYDINWYKKAASFTWAEVDISKDINDWNKLTDPDKYFIKH 61
QY 92 ILAFFAASDGINVENLVERSQEVQVPEARCFYGFQILLIENHSEMYSLLLDITYRDPKK 151

Db 62 VLAFFAASDGINVENLAERFCTEQITEARCFYGFQMAIENHSEMYSLLLDITYVKDSNE 121
QY 152 REFLFNAIETMPYKVKADWALRIADRKSTFGERVVFAAFAVEGFFSGSFAAIFWLKXR 211
Db 122 KNYLFNAIETMPCKVKKADWAQKWIHD-SAGYGERLIAFAAAGVIFGSGSFASIFWLKXR 180
QY 212 GLMPGLTFSNELSRDEGLHCDFACLMFQYLVNKPSEERVREIIVDAVKIEQEFLEALP 271
Db 181 GLMPGLTFSNELSRDEGLHCDFACLMFQYLVNKPSEERVRSIITDAVSIEQEFLEALP 240
QY 272 VGLIGNCILMKQYIEFVADRLVELGFSKVFQAEENPDMENISLEGKTNFFEKRVSEY 331
Db 241 VKLIGNCENMKTYIEFVADRLVELGFSKVFQAEENPDMENISLEGKTNFFEKRVSEY 300
QY 332 QRFVMAETTDNVFTLDADF 351
Db 301 QKGMVMSQ-EDNHFSLDVDF 319
RESULT 9
T28466
ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain - variola major virus
C;Species: variola major virus
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 15-Sep-2003
C;Accession: T28466
R;Massung, R.F.; Esposito, J.J.; Liu, L.I.; Qi, J.; Utterback, T.R.; Knight, J.C.; Aubin
Nature 366, 748-751, 1993
A;Title: Potential virulence determinants in terminal regions of variola smallpox virus
A;Reference number: Z20488; MUID:94088747; PMID:8264798
A;Accession: T28466
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-333 <MAS>
A;Cross-references: EMBL:L22579; NID:G623595; PID:AAA60776.1; PID:G438946
A;Experimental source: strain "Bangladesh-1975"
C;Function:
A;Description: catalyzes the reduction of ribonucleoside diphosphate to deoxyribonucleoside
C;Superfamily: ribonucleoside reductase small subunit
C;Keywords: deoxyribonucleotide biosynthesis; oxidoreductase
Query Match 70.2%; Score 1279; DB 2; Length 333;
Best Local Similarity 74.7%; Pred. No. 3.5e-97;
Matches 239; Conservative 35; Mismatches 44; Indels 2; Gaps 2;
QY 32 EPLLKSSRRFVFPPIQYDINWYKQAQASFTWAEVDLSKDLPHNKLKADKYPFISH 91
Db 16 EPLIAPNPNRFVFPPIQYDINWYKKAASFTWAEVDISKDINDWNKLTDPDKYFIKH 75
QY 92 ILAFFAASDGINVENLVERSQEVQVPEARCFYGFQILLIENHSEMYSLLLDITYRDPKK 151
Db 76 VLAFFAASDGINVENLAERFCTEQITEARCFYGFQMAIENHSEMYSLLLDITYVKDSNE 135
QY 152 REFLFNAIETMPYKVKADWALRIADRKSTFGERVVFAAFAVEGFFSGSFAAIFWLKXR 211
Db 136 KNYLFNAIETMPCKVKKADWAQKWIHD-SAGYGERLIAFAAAGVIFGSGSFASIFWLKXR 194
QY 212 GLMPGLTFSNELSRDEGLHCDFACLMFQYLVNKPSEERVREIIVDAVKIEQEFLEALP 271
Db 195 GLMPGLTFSNELSRDEGLHCDFACLMFQYLVNKPSEERVRSIITDAVSIEQEFLEALP 254
QY 272 VGLIGNCILMKQYIEFVADRLVELGFSKVFQAEENPDMENISLEGKTNFFEKRVSEY 331
Db 255 VKLIGNCENMKTYIEFVADRLVELGFSKVFQAEENPDMENISLEGKTNFFEKRVSEY 314
QY 332 QRFVMAETTDNVFTLDADF 351
Db 315 QKGMVMSQ-EDNHFSLDVDF 333
RESULT 10
H36839
C8L protein - variola virus (strain India-1967)

C:Species: variola virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C:Accession: H36839
R:Blinov, V.M.
submitted to GenBank, November 1992
A:Reference number: A36859
A:Accession: H36839
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-333 <BL>
A:Cross-references: UNIPROT:P33799; GB:X69198; NID:G456758; PIDN:CAA48969.1; PID:G297209
C:Superfamily: ribonucleoside reductase small subunit

Query Match 69.5%; Score 1266; DB 2; Length 333;
Best Local Similarity 74.1%; Pred. No. 4.1e-96;
Matches 237; Conservative 35; Mismatches 46; Indels 2; Gaps 2;

QY 32 EPLLKSRFRVIFPIQYDPDIWMYKQAQAFWTAEEVDLSKDPHWNKLKADKPYFISH 91
DB 16 EPLAKNPFRVIFPIQYDHIWMYKKAASFWTVEVDISKINDWNKLTDPKPYFIKH 75

QY 92 ILAPFAASDGIWNNLVHSEVQVPEARCFYGFQILLENHSEMYSLIDTYIRDPKK 151
DB 76 VLAPFAASDGIWNNLVHSEVQVPEARCFYGFQILLENHSEMYSLIDTYIRDPKK 135

QY 152 REFLNAIETMPVYKKAQADWLRWTRKSTGERVVAFAAEGVFFSGSPAAIFWLKXR 211
DB 136 KNYLFNAIETMPVYKKAQADWLRWTRKSTGERVVAFAAEGVFFSGSPAAIFWLKXR 194

QY 212 GLMPGLTFNLSRDEGLHCDPACILMFQYLVNKPSEERVREIIVDAVKIEQBFLEALP 271
DB 195 GLMPGLTFNLSRDEGLHCDPACILMFQYLVNKPSEERVREIIVDAVKIEQBFLEALP 254

QY 272 VGLIGNCILMKQYIEFVADRLLELGVSKVQFAENPDPFMENISLEGKTNFPEKRVSY 331
DB 255 VKLIGNCEMKYIEFVADRLLELGVSKVQFAENPDPFMENISLEGKTNFPEKRVSY 314

QY 332 QRFVMAETTDNVFTLDADF 351
DB 315 QKMGVMSQ-EDNHFSLDADF 333

RESULT 11
T30470
ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain - Lymantria dispar nuclea
N:Alternate names: ribonucleotide reductase small subunit homolog
C:Species: Lymantria dispar nuclear polyhedrosis virus, LdMNPV
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T30470
R:Kuzio, J.; Pearson, M.N.; Harwood, S.H.; Funk, C.J.; Evans, J.T.; Slavicek, J.M.; Rohr
Virology 253, 17-34, 1999
A:Title: Sequence and analysis of the genome of a baculovirus pathogenic for Lymantria d
A:Reference number: Z20836; MUID:99124785; PMID:9887315
A:Accession: T30470
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-348 <XU>
A:Cross-references: UNIPROT:Q9YMK7; EMBL:AF081810; PIDN:AAC70306.1
C:Function:
A:Description: catalyzes the reduction of ribonucleoside diphosphate to deoxyribonucleos
C:Superfamily: ribonucleoside reductase small subunit
C:Keywords: oxidoreductase; pyrimidine deoxynucleotide metabolism

Query Match 69.1%; Score 1258.5; DB 2; Length 348;
Best Local Similarity 70.1%; Pred. No. 1.8e-95;
Matches 246; Conservative 37; Mismatches 59; Indels 9; Gaps 3;

QY 1 MGDPRPEAAGLDQDERSSDTNESEIKSNBPLKSRFRVIFPIQYDPDIWMYKQAQ 60
DB 7 MPAPERP-----FDPRAPAP--ERFPDPRAEPLLRNRRFRVIFPIQYDPDIWMYKKA 58

QY 61 ASFWTAEEVDLSKDPHWNKLKADKPYFISHILAPFAASDGIWNNLVHSEVQVPEA 120

Db 59 ASFWTVEVDLSKDTSDWERNLNDNERHFIKHVLAPFAASDGIWNNLVHSEVQVPEA 118
QY 121 RCFYGFQILLENHSEMYSLIDTYIRDPKKRBFNAIETMPVYKKAQADWLRWTRK 180
Db 119 RCFYGFQIAMENHSEMYSLIDTYIRDSSEKRLNNAIETMPCKKAEWALWIAGRE 178
QY 181 STGERVVAFAAEGVFFSGSPAAIFWLKRGKGLMPGLTFNLSRDEGLHCDPACILMFQ 240
Db 179 AAFGERLVAFAAEGVFFSGSPAAIFWLKRGKGLMPGLTFNLSRDEGLHCDPACILMFQ 238
QY 241 YLVNKPSEERVREIIVDAVKIEQBFLEALPVGLIGNCILMKQYIEFVADRLLELGV 300
Db 239 HLQRFSAARVREIIVDAVKIEQBFLEALPVGLIGNCILMKQYIEFVADRLLELGV 298
QY 301 KVFQAEENPDPFMENISLEGKTNFPEKRVSEVQVFAVMAETTDNVFTLDADF 351
Db 299 KHYNTNPDLDFMNLISLEGKTNFPEKRVSEVQVFAVMAETTDNVFTLDADF 348

RESULT 12

RDS2R

ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain - Atlantic surf clam (frag
N:Alternate names: ribonucleotide reductase small chain
C:Species: Spisula solidissima (Atlantic surf clam)
C:Date: 28-Dec-1987 #sequence_revision 30-Sep-1988 #text_change 09-Jul-2004
C:Accession: A22259
R:Standart, N.M.; Bray, S.J.; George, E.L.; Hunt, T.; Ruderman, J.V.
J. Cell Biol. 100, 1968-1976, 1985
A:Title: The small subunit of ribonucleotide-diphosphate reductase is encoded by one of t
A:Reference number: A22259; MUID:85207963; PMID:2987274
A:Accession: A22259
A:Molecule type: DNA
A:Residues: 1-299 <STA>
A:Cross-references: UNIPROT:P07201
C:Comment: The synthesis of this polypeptide chain is triggered by fertilization of the e
C:Function:
A:Description: catalyzes the reduction of ribonucleoside diphosphate to deoxyribonucleos
C:Superfamily: ribonucleoside reductase small subunit
C:Keywords: deoxyribonucleotide biosynthesis; iron; metalloprotein; oxidoreductase
F;45,76,79,139,173,176/Binding site: 2Fe-O cluster (Asp, Glu, His, Glu, His) #statu
F;83/Active site: Tyr (stable tyrosyl radical) #status predicted

Query Match 66.8%; Score 1217; DB 1; Length 299;
Best Local Similarity 77.7%; Pred. No. 3.7e-92;
Matches 233; Conservative 33; Mismatches 30; Indels 4; Gaps 3;

QY 55 MYKQAQASFWTAEVDLSKDPHWNKLKADKPYFISHILAPFAASDGIWNNLVHSEVQ 114
Db 1 MYKKAASFW-AEVDLSKDMAHWSLKKEKFISHVLAPFAASDGIWNNLVHSEVQ 59

QY 115 VQVPEARCFYGFQILLENHSEMYSLIDTYIRDPKKRBFNAIETMPVYKKAQADWLR 174
Db 60 VGVTEARCFYGFQIAMENHSEMYSLIDTYIRDPKKRBFNAIETMPVYKKAQADWLR 119

QY 175 WIADRKSTGERVVAFAAEGVFFSGSPAAIFWLKRGKGLMPGLTFNLSRDEGLHCDP 234
Db 120 WINDSSSYAARVVAFAAEGVFFSGSPAAIFWLKRGKGLMPGLTFNLSRDEGLHCDP 179

QY 235 ACLMFQYLVNKPSEERVREIIVDAVKIEQBFLEALPVGLIGNCILMKQYIEFVADRL 294
Db 180 ACLMFQYLVNKPSEERVREIIVDAVKIEQBFLEALPVGLIGNCILMKQYIEFVADRL 239

QY 295 VELGFSKVFQAEENPDPFMENISLEGKTNFPEKRVSEVQVFAVMA--ETTD-NVFTLDADF 351
Db 240 LELKCDKLYNKENPDPFMENISLEGKTNFPEKRVSEVQVFAVMA--ETTD-NVFTLDADF 299

RESULT 13

T18876

hypotheical protein C03C10.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T18876

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2005, 06:36:26 ; Search time 71 Seconds
(without alignments)
2531.546 Million cell updates/sec

Title: US-10-698-228-1

Perfect score: 1821

Sequence: 1 MGDPRPEAAGLDQDERSSS.....QRFAVMAETTDNVTLADAF 351

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1821	100.0	351	2 Q7LG56	Q7LG56 homo sapien
2	1817	99.8	351	2 Q9NUW3	Q9NUW3 homo sapien
3	1725	94.7	351	2 Q6PEE3	Q6PEE3 mus musculus
4	1534	84.2	297	2 Q6A141	Q6A141 homo sapien
5	1514	83.1	299	2 Q75PY9	Q75PY9 homo sapien
6	1474	80.9	285	2 Q86VE3	Q86VE3 homo sapien
7	1454	79.8	389	1 R1R2_HUMAN	P31350 homo sapien
8	1443	79.2	390	1 R1R2_MOUSE	P11157 mus musculus
9	1439	79.0	386	2 Q6BEP1	Q6BEP1 xenopus tro
10	1431	78.6	386	2 Q801Q4	Q801Q4 xenopus lae
11	1427	78.4	386	2 Q8AVY2	Q8AVY2 xenopus lae
12	1425.5	78.3	386	1 R1R2_BRARE	P79733 brachydanio
13	1421	78.0	386	1 R1R2_MESAU	Q60561 mesocricetu
14	1420.5	78.0	386	2 Q6D144	Q6D144 brachydanio
15	1399.5	76.9	378	2 Q6P876	Q6P876 xenopus tro
16	1385	76.1	406	2 Q61P47	Q61P47 xenopus lae
17	1382.5	75.9	388	2 Q7PF28	Q7PF28 anopheles g
18	1382.5	75.9	426	2 Q7O1F4	Q7O1F4 anopheles g
19	1370	75.2	349	2 Q7Z1W0	Q7Z1W0 brachydanio
20	1358	74.6	397	2 Q95VP8	Q95VP8 aedes aegypt
21	1349.5	74.1	399	2 Q27124	Q27124 urechis cau
22	1348	74.0	393	1 R1R2_DROME	P48592 drosophila
23	1335	73.3	398	2 Q9XYN8	Q9XYN8 aedes albop
24	1331	73.1	384	1 R1R2_SPISO	P07201 spissula sol
25	1284	70.5	403	2 Q6CFU6	Q6CFU6 yarowia li
26	1280	70.3	319	2 Q57175	Q57175 vaccinia vi
27	1280	70.3	319	2 Q76Q46	Q76Q46 variola min
28	1280	70.3	319	2 Q89087	Q89087 variola vir
29	1279	70.2	319	1 R1R2_VACCV	P11158 vaccinia vi
30	1279	70.2	319	2 Q76ZK1	Q76ZK1 vaccinia vi
31	1277	70.1	319	2 Q6RZQ9	Q6RZQ9 rabbitpox v

P20493 vaccinia vi
Q81lh6 ectromelia
Q89559 variola vir
Q8gm14 cowpox viru
P29883 vaccinia vi
Q8V544 monkeypox v
P87632 cowpox viru
Q8V2x8 camelpox vi
Q775z5 camelpox vi
Q91fe9 vaccinia vi
P33799 variola vir
Q9C167 neurospora
Q9pxr4 variola vir
Q9ymk7 lymantria d

32 1276 70.1 319 1 R1R2_VACCC
33 1276 70.1 319 2 Q8JLH6
34 1276 70.1 319 2 Q89559
35 1274 70.0 333 2 Q8QNI4
36 1273 69.9 319 1 R1R2_VACCP
37 1271 69.8 319 2 Q8V544
38 1270 69.7 319 2 P87632
39 1268 69.6 319 2 Q8V2x8
40 1268 69.6 319 2 Q775Z5
41 1267 69.6 319 2 Q9JPE9
42 1266 69.5 319 1 R1R2_VARV
43 1264.5 69.4 410 1 R1R2_NEUCR
44 1260 69.2 333 2 Q9PXR4
45 1258.5 69.1 348 2 Q9YMK7

ALIGNMENTS

RESULT 1

Q7LG56 PRELIMINARY; PRT; 351 AA.
AC Q7LG56; Q9NPD6; Q9NTD8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Ribonucleotide reductase (Hypothetical protein DKFZp761E1312) (P53-
DE Inducible ribonucleotide reductase small subunit 2).
GN Name=p53R2; Synonyms=DKFZp761E1312;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=20179179; PubMed=10716435; DOI=10.1038/35003506;
RA Tanaka H., Arakawa H., Yamaguchi T., Shiraishi K., Fukuda S.,
RA Matsui K., Takei Y., Nakamura Y.;
RT "A ribonucleotide reductase gene involved in a p53-dependent cell-
cycle checkpoint for DNA damage.";
RL Nature 404:42-49 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Amgdales;
RG The German CDNA Consortium;
RA Ansoerge W., Krieger S., Regiert T., Rittmueller C., Schwager B.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Ugai H., Yokoyama K.K.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB036532; BAA92493.1; -.
DR EMBL; AB036524; BAA92493.1; JOINED.
DR EMBL; AB036525; BAA92493.1; JOINED.
DR EMBL; AB036526; BAA92493.1; JOINED.
DR EMBL; AB036527; BAA92493.1; JOINED.
DR EMBL; AB036528; BAA92493.1; JOINED.
DR EMBL; AB036529; BAA92493.1; JOINED.
DR EMBL; AB036530; BAA92493.1; JOINED.
DR EMBL; AB036531; BAA92493.1; JOINED.
DR EMBL; AL137348; CAB70703.2; -.
DR EMBL; AB166671; BAD12267.1; -.
DR EMBL; AB036063; BAA92434.1; -.
DR Genew; HGNC:17296; RRM2B.
DR GO; GO:0004748; P:ribonucleoside-diphosphate reductase activity; IEA.
DR GO; GO:0009186; P:deoxyribonucleoside diphosphate metabolism; IEA.
DR InterPro; IPR009078; Ferritin/RR like.
DR InterPro; IPR000358; Ribonucleotide reductase.
DR Pfam; PF00268; Ribonucleotide reductase.
DR PROSITE; PS00368; RIBONUC RED SM; 1.
KW Hypothetical protein.
SQ SEQUENCE 351 AA; 40736 MW; 6D008687E6F40994 CRC64;

Query Match 100.0%; Score 1821; DB 2; Length 351;
Best Local Similarity 100.0%; Pred. No. 6, 1e-138;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVFPFIQYDPDIWKMYKQAO 60
DB 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVFPFIQYDPDIWKMYKQAO 60
QY 61 ASFWTAAEEVDLSKDLPHNKLKADKEYFISHILAFPAASDGIWNLVERFSQEVQVPEA 120
DB 61 ASFWTAAEEVDLSKDLPHNKLKADKEYFISHILAFPAASDGIWNLVERFSQEVQVPEA 120
QY 121 RCFYGFQILLENVHSEMYSLIDTYIRDPKKEFLFNAIETMPYVKKKADWALRWIADRK 180
DB 121 RCFYGFQILLENVHSEMYSLIDTYIRDPKKEFLFNAIETMPYVKKKADWALRWIADRK 180
QY 181 STFGERVVAFPAAVEGVFFSGSFAAIFWLKRGMLMPLGTFNSNELISRDEGLHCDFACLMFQ 240
DB 181 STFGERVVAFPAAVEGVFFSGSFAAIFWLKRGMLMPLGTFNSNELISRDEGLHCDFACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLVLVGLGFS 300
DB 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLVLVGLGFS 300
QY 301 KVFQAEPPDFMENISLEGKTNFFKRVSEYQRFVAVMAETTDNVFTLDADF 351
DB 301 KVFQAEPPDFMENISLEGKTNFFKRVSEYQRFVAVMAETTDNVFTLDADF 351

RESULT 2
Q9NUM3 PRELIMINARY; PRT; 351 AA.

AC Q9NUM3
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Hypothetical protein FLJ11103.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta; DOI=10.1038/ng1285;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
RA Abe K., Kimihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hiraoa S., Chiba Y., Ishida S.,
RA Oono T., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura J.,
RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
RA Musahiro K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsumura H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiya H., Satch N., Takami S., Terashima Y., Suzuki O.,
RA Nishigaki S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takenoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Omori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,

RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs";
RL Nat. Genet. 36:40-45 (2004).
DR ENBL; AK001965; BAA92005.1; --
DR HSSP; P11157; 1XSM.
DR GO; GO:0004748; P:ribonucleoside-diphosphate reductase activity; IEA.
DR GO; GO:0009186; P:deoxyribonucleoside diphosphate metabolism; IEA.
DR InterPro; IPR009078; Ferritin/RR like.
DR InterPro; IPR000358; Ribonucleotidase.
DR Pfam; PF00268; Ribonucleotidase; 1.
DR PROSITE; PS00368; RIBED5 SMALL; 1.
SQ SEQUENCE 351 AA; 47074 MW; 6D009B2D59E9A323 CRC64;

Query Match 99.8%; Score 1817; DB 2; Length 351;
Best Local Similarity 99.7%; Pred. No. 1, 3e-137;
Matches 350; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVFPFIQYDPDIWKMYKQAO 60
DB 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVFPFIQYDPDIWKMYKQAO 60
QY 61 ASFWTAAEEVDLSKDLPHNKLKADKEYFISHILAFPAASDGIWNLVERFSQEVQVPEA 120
DB 61 ASFWTAAEEVDLSKDLPHNKLKADKEYFISHILAFPAASDGIWNLVERFSQEVQVPEA 120
QY 121 RCFYGFQILLENVHSEMYSLIDTYIRDPKKEFLFNAIETMPYVKKKADWALRWIADRK 180
DB 121 RCFYGFQILLENVHSEMYSLIDTYIRDPKKEFLFNAIETMPYVKKKADWALRWIADRK 180
QY 181 STFGERVVAFPAAVEGVFFSGSFAAIFWLKRGMLMPLGTFNSNELISRDEGLHCDFACLMFQ 240
DB 181 STFGERVVAFPAAVEGVFFSGSFAAIFWLKRGMLMPLGTFNSNELISRDEGLHCDFACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLVLVGLGFS 300
DB 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLVLVGLGFS 300
QY 301 KVFQAEPPDFMENISLEGKTNFFKRVSEYQRFVAVMAETTDNVFTLDADF 351
DB 301 KVFQAEPPDFMENISLEGKTNFFKRVSEYQRFVAVMAETTDNVFTLDADF 351

RESULT 3
Q6PEE3 PRELIMINARY; PRT; 351 AA.

AC Q6PEE3
DT 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
DE Ribonucleotide reductase M2 B (P53 inducible).
GN Name=Rrm2b;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Srausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RT Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Strausberg R.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC058103; AAH58103.1; -
DR GO; GO:0004748; F:ribonucleoside-diphosphate reductase activity; IDA.
DR GO; GO:0009263; P:deoxyribonucleotide biosynthesis; IDA.
DR InterPro; IPR009078; Ferritin/RR like.
DR InterPro; IPR000358; Ribonuc. redctse.
DR Pfam; PF00268; Ribonuc. red sm; 1.
DR PROSITE; PS00368; RIBORED_SMALL; 1.
SQ SEQUENCE 351 AA; 40803 MW; 4E1259233C9C8A9 CRC64;

Query Match 94.7%; Score 1725; DB 2; Length 351;
Best Local Similarity 93.4%; Pred. No. 3.2e-130;
Matches 328; Conservative 14; Mismatches 9; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDQDERSSTNESEIKSNBPLRKSSRRVFPIQVDPDIWKYKQQAQ 60
DB 1 MGDPERPEAARPEKEQGLCSTEEVNSBPLRKSSRRVFPIQVDPDIWKYKQQAQ 60

QY 61 ASFWTAEVDLSKDLPHWNKLKADKEYFISHILAFFAASDGIWNLVERFSQEVPEA 120
DB 61 ASFWTAEVDLSKDLPHWNKLKADKEYFISHILAFFAASDGIWNLVERFSQEVPEA 120

QY 121 RCFYGFQIILNHNHSEMSYLLIDTYIRDPKREFLNAETMPYVKKADWALRIADRK 180
DB 121 RCFYGFQIILNHNHSEMSYLLIDTYIRDPKREFLNAETMPYVKKADWALRIADRK 180

QY 181 STFGERVVAFVAVGPFSSGFAAIFWLKRGMLPGLTFSENLISRDGLHCDPACLMFQ 240
DB 181 STFGERVVAFVAVGPFSSGFAAIFWLKRGMLPGLTFSENLISRDGLHCDPACLMFQ 240

QY 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLV 300
DB 241 YLVNKPSEDRVREIIVDAVQIEQEFLEALPVGLIGMNCILMKQYIEFVADRLV 300

QY 301 KVFQAEVDFPMENISLEGKTNFPEKRVSEYQRFVAVMAETTDNVFTLDADF 351
DB 301 KIFQAEVDFPMENISLEGKTNFPEKRVSEYQRFVAVMAETTDNVFTLDADF 351

RESULT 4
Q6A141 PRELIMINARY; PRT; 297 AA.
AC Q6A141
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Hypothetical protein DKFZp686M05248.
GN Name=DKFZp686M05248;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RG The German cDNA Consortium;
RA Ottenwelder B., Obermaier B., Deutschenbaur S., Schaiipp A.,
RA Mewes H.W., Weil B., Anid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR627376; CAH10473.1; -
DR GO; GO:0004748; F:ribonucleoside-diphosphate reductase activity; IEA.
DR GO; GO:0009186; P:deoxyribonucleoside diphosphate metabolism; IEA.

DR InterPro; IPR009078; Ferritin/RR like.
DR InterPro; IPR000358; Ribonuc. redctse.
DR Pfam; PF00268; Ribonuc. red sm; 1.
DR PROSITE; PS00368; RIBORED_SMALL; 1.
KW Hypothetical protein.
SQ SEQUENCE 297 AA; 34498 MW; 2BE522F2CC158A02 CRC64;

Query Match 84.2%; Score 1534; DB 2; Length 297;
Best Local Similarity 99.7%; Pred. No. 5.7e-115;
Matches 296; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 55 MYKQAQASFWTAEVDLSKDLPHWNKLKADKEYFISHILAFFAASDGIWNLVERFSQ 114
DB 1 MYKQAQASFWTAEVDLSKDLPHWNKLKADKEYFISHILAFFAASDGIWNLVERFSQ 60

QY 115 VQVPEARCFYGFQIILNHNHSEMSYLLIDTYIRDPKREFLNAETMPYVKKADWALR 174
DB 61 VQVPEARCFYGFQIILNHNHSEMSYLLIDTYIRDPKREFLNAETMPYVKKADWALR 120

QY 175 WTADKSTFGERVVAFVAVGPFSSGFAAIFWLKRGMLPGLTFSENLISRDGLHCD 234
DB 121 WTADKSTFGERVVAFVAVGPFSSGFAAIFWLKRGMLPGLTFSENLISRDGLHCD 180

QY 235 ACLMFOYLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRL 294
DB 181 ACLMFOYLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRL 240

QY 295 VELGFSKVFOAENPDPFMENISLEGKTNFPEKRVSEYQRFVAVMAETTDNVFTLDADF 351
DB 241 VELGFSKVFOAENPDPFMENISLEGKTNFPEKRVSEYQRFVAVMAETTDNVFTLDADF 297

RESULT 5
Q75PY9 PRELIMINARY; PRT; 299 AA.
ID Q75PY9
AC Q75PY9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE P53-inducible ribonucleotide reductase small subunit 2 long form.
GN Name=P53R2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ugai H., Yokoyama K.K.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB163437; BAD11774.1; -
DR GO; GO:0004748; F:ribonucleoside-diphosphate reductase activity; IEA.
DR GO; GO:0009186; P:deoxyribonucleoside diphosphate metabolism; IEA.
DR InterPro; IPR009078; Ferritin/RR like.
DR InterPro; IPR000358; Ribonuc. redctse.
DR Pfam; PF00268; Ribonuc. red sm; 1.
DR PROSITE; PS00368; RIBORED_SMALL; 1.
SQ SEQUENCE 299 AA; 34528 MW; 4705C44389EB689B CRC64;

Query Match 83.1%; Score 1514; DB 2; Length 299;
Best Local Similarity 85.2%; Pred. No. 2.3e-113;
Matches 299; Conservative 0; Mismatches 0; Indels 52; Gaps 1;

QY 1 MGDPERPEAAGLDQDERSSTNESEIKSNBPLRKSSRRVFPIQVDPDIWKYKQQAQ 60
DB 1 MGDPERPEAAGLDQD-----

QY 61 ASFWTAEVDLSKDLPHWNKLKADKEYFISHILAFFAASDGIWNLVERFSQEVPEA 120
DB 16 -----EVDLSKDLPHWNKLKADKEYFISHILAFFAASDGIWNLVERFSQEVPEA 68

QY 121 RCFYGFQIILNHNHSEMSYLLIDTYIRDPKREFLNAETMPYVKKADWALRIADRK 180
DB 69 RCFYGFQIILNHNHSEMSYLLIDTYIRDPKREFLNAETMPYVKKADWALRIADRK 128

DR GO: 0009262; P-deoxyribonucleotide metabolism; IDA.
 DR InterPro: IPR000978; Ferritin/RR like.
 DR InterPro: IPR000358; Ribonuc redctse.
 DR Pfam: PF00268; Ribonuc_red_sm; 1.
 DR PROSITE: PS00368; RIBORED_SMALL; 1.
 DR 3D-structure; DNA replication; Iron; Metal-binding; Oxidoreductase.
 KW METAL 139 139 Iron 1.
 FT METAL 170 170 Iron 1 and 2.
 FT METAL 173 173 Iron 1.
 FT METAL 233 233 Iron 2.
 FT METAL 267 267 Iron 1 and 2.
 FT METAL 270 270 Iron 2.
 FT ACT_SITE 177 177 By similarity.
 FT HELIX 68 70
 FT TURN 72 74
 FT HELIX 89 99
 FT TURN 100 101
 FT HELIX 105 107
 FT TURN 112 113
 FT HELIX 114 118
 FT TURN 119 119
 FT HELIX 122 148
 FT TURN 149 149
 FT HELIX 150 153
 FT HELIX 157 184
 FT HELIX 188 195
 FT TURN 196 196
 FT HELIX 197 200
 FT HELIX 202 204
 FT HELIX 205 216
 FT TURN 218 219
 FT HELIX 222 235
 FT TURN 236 237
 FT HELIX 238 249
 FT TURN 250 251
 FT HELIX 254 279
 FT TURN 280 280
 FT HELIX 287 306
 FT TURN 307 308
 FT HELIX 311 314
 FT TURN 315 315
 FT HELIX 318 335
 FT TURN 336 337
 FT HELIX 349 351
 SQ SEQUENCE 390 AA; 45095 MW; AC7ACC4PAP8A4A2P CRC64;
 Query Match 79.2%; Score 1443; DB 1; Length 390;
 Best Local Similarity 78.9%; Pred. No. 1.6e-107;
 Matches 269; Conservative 40; Mismatches 26; Indels 6; Gaps 1;
 QY 17 RSSDTSNESEIKS-----NEEPLLRKSSRRFVFIPIQYDPIWKMVKQAASFTWTAEEVD 70
 DB 50 RIFQDSAELESKAPTNSVEDEPLLRNRRFVFIPIQYDPIWKMVKQAASFTWTAEEVD 109
 QY 71 LSKDLPHNKLKADKEYFISHILAPFAASDGIWNLVERFSEVQVPEARCFYGFQILI 130
 DB 110 LSKDLQHWALKPDERHFIHVLAPFAASDGIWNLVERFSEVQVTEARCFYGFQIAM 169
 QY 131 ENVHSEMYSLLDITYIRDPKREFLFNAIETMPYVKKADWALRWADRSKTFGERVAF 190
 DB 170 ENHSEMYSLLDITYIKDKPKREYLFNAIETMPYVKKADWALRWADRSKTFGERVAF 229
 QY 191 AAVEGVFFSGSFAAIFWLKRGGLMPGLTFSNELISRDEGLHCDPACLMFYLVNKPSE 250
 DB 230 AAVEGIFFGSGSFASIFWLKRGGLMPGLTFSNELISRDEGLHCDPACLMFYLVNKPSE 289
 QY 251 VREIIVDAVKIQEFLTEALPVGLIGMNCILMKQYIEFVADRLLLVELGFSKVFQAEV 310
 DB 290 VREITINAVRIQSEFLTEALPVKLGIMNCILMKQYIEFVADRLLLVELGFSKVFQAEV 349
 QY 311 FMENISLEKTNFFPKRVSEYQRFVAAETATDNVFTLDADF 351
 DB 350 FMENISLEKTNFFPKRVSEYQRFVAAETATDNVFTLDADF 390

RESULT 9

Q68EP1 PRELIMINARY; PRT; 386 AA.
 ID Q68EP1
 AC Q68EP1
 DT 25-OCT-2004 (TREMBlrel. 28, Created)
 DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
 DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
 DE Hypothetical protein.
 OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OC NCBI_TaxID=8364;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Matek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences".
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RA Klein S., Gerhard D.S.;
 RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC080161; AAH80161.1; -
 DR InterPro: IPR009078; Ferritin/RR like.
 DR InterPro: IPR000358; Ribonuc redctse.
 DR Pfam: PF00268; Ribonuc_red_sm; 1.
 DR PROSITE: PS00368; RIBORED_SMALL; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 386 AA; 44540 MW; 277DF859B6053824 CRC64;
 Query Match 79.0%; Score 1439; DB 2; Length 386;
 Best Local Similarity 78.1%; Pred. No. 3.4e-107;
 Matches 268; Conservative 39; Mismatches 36; Indels 0; Gaps 0;
 QY 9 AAGLDQDERSSDTSNESEIKSNEEPLLRKSSRRFVFIPIQYDPIWKMVKQAASFTWTAEE 68
 DB 44 ARNIFQEAETTKSKAPKDPRIQDEPLLRNRRFVFIPIQYDPIWKMVKQAASFTWTAEE 103
 QY 69 VDLSKDLPHNKLKADKEYFISHILAPFAASDGIWNLVERFSEVQVPEARCFYGFQI 128
 DB 104 VDLSKDLRHVESLKAEKEYFISHVLAPFAASDGIWNLVERFSEVQVTEARCFYGFQI 163
 QY 129 LIENHSEMYSLLDITYIRDPKREFLFNAIETMPYVKKADWALRWADRSKTFGERV 188
 DB 164 AMENIHSEMYSLLDITYIKDKPKREYLFNAIETLPYVKKADWALRWADRSKTFGERV 223
 QY 189 AFAAVEGVFFSGSFAAIFWLKRGGLMPGLTFSNELISRDEGLHCDPACLMFYLVNKPSE 248
 DB 224 AFAAVEGIFFGSGSFASIFWLKRGGLMPGLTFSNELISRDEGLHCDPACLMFYLVNKPSE 283
 QY 249 ERVREIIVDAVKIQEFLTEALPVGLIGMNCILMKQYIEFVADRLLLVELGFSKVFQAEV 308

```

Db 284 ERVVQITDAVQIEQEFLEALPVNIGMNCITLKMQVIEFVADRLLELGFNKFVFRATNP 343
Qy 309 FDMENISLEGKTNFFKRVSEYQRFVAVMAETTDNVFTLDADF 351
Db 344 FDMENISLEGKTNFFKRVSEYQRFVAVMAETTDNVFTLDADF 386

RESULT 10
Q80104
ID Q80104 PRELIMINARY; PRT; 386 AA.
AC Q80104;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Rm2-prov protein
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC047975; AAH47975.1; --
DR HSSP; P11157; 1XSM.
DR GO; GO:0004748; Fribonucleoside-diphosphate reductase activity; IEA.
DR GO; GO:0009186; P-deoxyribonucleoside diphosphate metabolism; IEA.
DR InterPro; IPR009078; Ferritin/RR like.
DR InterPro; IPR000358; Ribonucleotide.
DR Pfam; PF00268; Ribonuc red sm; 1.
DR PROSITE; PS00368; RIBRED_SWALL; 1.
DR SEQUENCE 386 AA; 44595 MW; 3C12900A1CA7D54A CRC64;

Query Match 78.6%; Score 1431; DB 2; Length 386;
Best Local Similarity 79.5%; Pred. No. 1.5e-106;
Matches 267; Conservative 38; Mismatches 27; Indels 4; Gaps 1;

```

```

RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC041209; AAH41209.1; -.
DR HSSP; P11157; IXSM.
DR GO; GO:0004748; F:Ribonucleoside-diphosphate reductase activity; IEA.
DR GO; GO:000186; P:deoxyribonucleoside diphosphate metabolism; IEA.
DR InterPro; IPR009078; Ferritin/RR like.
DR InterPro; IPR000358; RibonucL reductase.
DR Pfam; PF00268; RibonucL reductase.
DR Pfam; PF00268; RibonucL reductase.
DR PROSITE; PS00368; RIBORED_SMALL; 1.
DR SEQUENCE 386 AA; 44635 MW; 4635FCC9BB18EA32 CRC64;

Query Match 78.4%; Score 1427; DB 2; Length 386;
Best Local Similarity 79.2%; Pred. No. 3.1e-106;
Matches 266; Conservative 36; Mismatches 30; Indels 4; Gaps 1;

QY 20 SDTNSEIKSN-----EPELLKSSRRFVPIQYDPDIWKMYKQAQSFMTAEVDLSKOL 75
Db 51 TETPKSKAPKNRLEDEPLKDKNPHRFVPIQYHDIWQYKKAASFMTAEVDLSKOL 110
QY 76 PHWNKLKADKPYFISHILAFFAASDGINVENLVERFSQVQVPEARCFYGFQILLENVHS 135
Db 111 QHWESLKKEEKFYFISHVLFAFASDGINVENLVERFSQVQVTEARCFYGFQIAMENIHS 170
QY 136 EMYSLLIDITYIRDPKREFLFNAIETMPYVKKADWALRWIADRKSTFGERVVAFPAVEG 195
Db 171 EMYSLLIDITYIRDPKREFLFNAIETPCVKKADWALRWIADRKSTFGERVVAFPAVEG 230
QY 196 VFFSGSFAAIFWLKRGKGLMPGLTFSNELISRDEGLHCDFACLMFOYLNVKPKSERVREII 255
Db 231 IFFSGSFASIFWLKRGKGLMPGLTFSNELISRDEGLHCDFACLMFKHLRKPSSERVVELI 290
QY 256 VDAVKIEQFLEALPVGLIGMNCILMKQYIEFVADRLLELVELGFSKVFQAEVNFPMENI 315
Db 291 TDAVQIEQFLEALPVGLIGMNCILMKQYIEFVADRLLELVELGFSKVFQAEVNFPMENI 350
QY 316 SLEGTNFFPEKRVSEYQRFVAVMAETDNTVTLDAF 351
Db 351 SLEGTNFFPEKRVSEYQRFVAVMAETDNTVTLDAF 386

RESULT 12
RIR2_BRARE STANDARD; PRT; 386 AA.
AC P79733;
DT 15-DEC-1998 (Rel. 37, Created)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Ribonucleoside-diphosphate reductase M2 chain (EC 1.17.4.1)
DE (Ribonucleotide reductase protein R2 class 1).
GN Name=rnm2; Synonyms=r2;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OC NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97137859; PubMed=8983196;
RA Mathews C.Z., Sjoeborg B.-M., Karlsson M.;
RT "Cloning and sequencing of cDNAs encoding ribonucleotide reductase
RT from zebrafish Danio rerio."
RL Mol. Mar. Biol. Biotechnol. 5:284-287(1996).
CC -!- FUNCTION: Provides the precursors necessary for DNA synthesis.
CC -!- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate +
CC thiorodoxin disulfide + H2O = ribonucleoside diphosphate +
CC thiorodoxin.
CC -!- COFACTOR: Binds 2 iron ions per subunit (By similarity).
CC -!- PATHWAY: DNA replication pathway; first step.
CC -!- SUBUNIT: Heterodimer of a large (M1) and a small chain (M2).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the ribonucleoside diphosphate reductase

```

```

small chain family.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; U57965; AAB37103.1; -.
DR ZFIN; ZDB-GENE-990415-25; rnm2.
DR InterPro; IPR009078; Ferritin/RR like.
DR InterPro; IPR000358; RibonucL reductase.
DR Pfam; PF00268; RibonucL reductase.
DR PROSITE; PS00368; RIBORED_SMALL; 1.
DR DNA replication; Iron; Metal-binding; Oxidoreductase.
FT METAL 135 135 Iron 1 (By similarity).
FT METAL 166 166 Iron 1 and 2 (By similarity).
FT METAL 169 169 Iron 1 (By similarity).
FT METAL 229 229 Iron 2 (By similarity).
FT METAL 263 263 Iron 2 (By similarity).
FT METAL 266 266 Iron 2 (By similarity).
FT ACT SITE 173 173 By similarity.
SQ SEQUENCE 386 AA; 44594 MW; C168846FB57F9F4E CRC64;

Query Match 78.3%; Score 1425.5; DB 1; Length 386;
Best Local Similarity 78.8%; Pred. No. 4.1e-106;
Matches 268; Conservative 36; Mismatches 31; Indels 5; Gaps 1;

QY 17 RSSSDTNSEIKS-----NEEPLKSSRRFVPIQYDPDIWKMYKQAQSFMTAEVDL 71
Db 47 RK1FDESEGSKAKKGAEEELPKENRHPVPIQYHDIWQYKKAASFMTAEVDL 106
QY 72 SKDLPHWNKLKADKPYFISHILAFFAASDGINVENLVERFSQVQVPEARCFYGFQILIE 131
Db 107 SKDLQHWDSLKDEERYFISHVLFAFASDGINVENLVERFSQVQVTEARCFYGFQIAME 166
QY 132 NVHSEMYSLIDITYIRDPKREFLFNAIETMPYVKKADWALRWIADRKSTFGERVVAF 191
Db 167 NIHSEMYSLIDITYIRDPKREFLFNAIETMPYVKKADWALRWIADRKSTFGERVVAF 226
QY 192 AVEGVFSGSFAAIFWLKRGKGLMPGLTFSNELISRDEGLHCDFACLMFOYLNVKPKSERV 251
Db 227 AVEGVFSGSFAAIFWLKRGKGLMPGLTFSNELISRDEGLHCDFACLMFKHLRKPSSETV 286
QY 252 REIIVDAVKIEQFLEALPVGLIGMNCILMKQYIEFVADRLLELVELGFSKVFQAEVNF 311
Db 287 KKIIMNAVEIQEFLTDALPVKLGIMNCILMKQYIEFVADRLLELVELGFSKVFQAEVNF 346
QY 312 MENISLEGKTNFFPEKRVSEYQRFVAVMAETDNTVTLDAF 351
Db 347 MENISLEGKTNFFPEKRVSEYQRFVAVMAETDNTVTLDAF 386

RESULT 13
RIR2_MESAU STANDARD; PRT; 386 AA.
AC Q60561;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Ribonucleoside-diphosphate reductase M2 chain (EC 1.17.4.1)
DE (Ribonucleotide reductase small chain).
GN Name=rnm2;
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Rodentia; Muridae; Cricetinae;
OC Mesocricetus.
OC NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.

```

359 KTNFFEKRVGEYQRMGVMS-----NSFTLDADF 386

RESULT 14

Q6DI44 PRELIMINARY; PRT; 386 AA.

ID Q6DI44

AC Q6DI44;

DT 25-OCT-2004 (TtEMBLrel. 28, Created)

DT 25-OCT-2004 (TtEMBLrel. 28, Last sequence update)

DT 25-OCT-2004 (TtEMBLrel. 28, Last annotation update)

DE Ribonucleotide reductase M2 polypeptide.

GN Names:rm2;

OS Brachydanio rerio (zebrafish) (Danio rerio).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;

OC Cyprinidae; Danio.

OX NCBI_TaxID=7955;

RN [1]

RN SEQUENCE FROM N.A.

RP TTSUVE=Embryo;

RC MEDLINE=223388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;

RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,

RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Uudin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Pahay J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

RA Krzywiński M.I., Skalka U., Smalhus D.E., Schnerch A., Schein J.E.,

RA Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences.";

RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RL [2]

RL SEQUENCE FROM N.A.

RP TTSUVE=Embryo;

RC Strausberg R.;

RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC075746; AAH75746.1; -

DR GO; GO:0004786; F:riboynucleoside-diphosphate reductase activity; IEA.

DR GO; GO:0009186; P:deoxyribonucleoside diphosphate metabolism; IEA.

DR InterPro; IPRO09078; Ferritin/RR like.

DR InterPro; IPRO00358; Ribonuclease.

DR Pfam; PF00268; Ribonuc_red_sm; 1.

DR PROSITE; PS00368; RIBOSED_SMALL; 1.

SQ SEQUENCE 386 AA; 44567 MW; 2CCACD4424DD1566 CRC64;

Query Match 78.0%; Score 1420.5; DB 2; Length 386;

Best Local Similarity 78.5%; Pred. No. 1e-105;

Matches 267; Conservative 37; Mismatches 31; Indels 5; Gaps 1;

QY 17 RSSSDTNESEIKS-----NEEPLRKSSRRFVFPIQYDPDIWMYKQAOASFWTAEVDL 71

DB 47 RKTFDESEGSKAKGAVEEPELLKNPHRFVFPIQYHDIWQYKAEASFMTAEVDL 106

QY 72 SKDLPTWNLKADKEYFISHILAFPAASDGIWNVLENVERFSQVQPEARCFYGFQILLIE 131

DB 107 SKDLQWDSLKDBERYFISHVLFAAASDGIWNVLENVERFTQEVQTEARCFYGFQIAE 166

QY 132 NVHSEMYSLLDITYIDPKKREFLNAITMPVYKKADWALRWIADRKSTFGERVVAFA 191

DB 167 NVHSEMYSLLDITYIDKSKREFLNAITMPVYKKADWALRWIADRKSTFGERVVAFA 226

QY 192 AVEGVFFSGSFAAIFWLKKRGLMPLGTLTFSNELISRDGLHCDFACLWFQYLVNKPSEERV 251

```
Db 227 AVEGIFFGSFASIFWLKKRGLMPLGTFNSNELISRDEGLHCDPACLMFKHLISKPSSETV 286
QY 252 REIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLVLELGFSGVFOAENPDF 311
Db 287 KKLIMNAVEIEQEFLEALPVGLIGMNCILMKQYIEFVADRLVLELGFSGVFOAENPDF 346
QY 312 MENISLEGKTNFPEKRVSEYQREFAVMAETTDNVFTLDADF 351
Db 347 MENISLEGKTNFPEKRVSEYQREFAVMAETTDNVFTLDADF 386

RESULT 15
Q6P876
ID AC Q6P876 PRELIMINARY; PRT; 378 AA.
DC Q6P876;
DT 05-JUL-2004 (T-EMBLrel. 27, Created)
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DE Hypothetical protein MGC75900.
GN Name=MGC75900;
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8364;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Sapleten L., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Heiton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalusz D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences."
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Gerhard D.S.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC061353.1; -
DR GO; GO:0004748; F:ribonucleoside-diphosphate reductase activity; IEA.
DR GO; GO:0009186; P:deoxyribonucleoside diphosphate metabolism; IEA.
DR InterPro; IPR009078; Ferritin/RR like.
DR InterPro; IPR000358; Ribonuclease.
DR Pfam; PF00268; Ribonuclease.
DR PROSITE; PS00368; RIBONUC REASE.
KW Hypothetical protein.
SQ SEQUENCE 378 AA; 43939 MW; 538D477FBB7CDB65 CRC64;
```

Search completed: October 30, 2005, 06:57:37
Job time : 72 secs

```
Query Match 76.9%; Score 1399.5; DB 2; Length 378;
Best Local Similarity 76.6%; Pred. No. 4.9e-104;
Matches 262; Conservative 45; Mismatches 28; Indels 7; Gaps 1;

QY 17 RSSSDNESEI-----KSNEPLRKSSRRVFIPIQYVDIWKYKQAQASFTAAEEV 69
Db 37 RTSRGTHKQWCSLKDPRIQDEPLLRDNPGRFVILPIEYHDIWQYKKAASFTAAEEV 96
QY 70 DLKSDLPHNKLKADKDYFISHILAFPAASDGIWNLVERFSQEVQVPCRCFYGFQIL 129
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2005, 06:50:41 ; Search time 25 Seconds
(without alignments)
1048.073 Million cell updates/sec

Title: US-10-698-228-1
Perfect score: 1821
Sequence: 1 MGDPERPEAGLDQDERSSS.....QRFAVMAETDNTVFTLDADF 351

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1821	100.0	351	4	US-10-019-733-1
2	1454	79.8	389	4	US-10-019-733-5
3	1454	79.8	389	4	US-09-949-016-6016
4	1454	79.8	453	4	US-09-949-016-7896
5	1131	62.1	411	4	US-09-248-796A-18017
6	1072.5	58.9	274	1	US-08-307-499-19
7	1072.5	58.9	274	3	US-09-299-268-19
8	974.5	53.5	322	1	US-08-136-743B-2
9	969.5	53.2	360	4	US-09-248-796A-18018
10	768	42.2	147	3	US-08-908-223-358
11	599	32.9	149	4	US-09-270-767-46236
12	495	27.2	323	4	US-09-902-540-10006
13	381.5	21.0	445	4	US-09-328-352-4714
14	377.5	20.7	476	4	US-09-252-991A-19358
15	335	18.4	365	4	US-09-198-452A-1059
16	335	18.4	380	4	US-09-438-185A-986
17	298	16.4	314	3	US-09-080-044-3
18	298	16.4	314	3	US-09-531-857A-3
19	294	16.1	56	3	US-08-905-223-362
20	220	12.1	382	4	US-09-543-681A-4239
21	203.5	11.2	412	4	US-09-488-039A-8180
22	201.5	11.1	401	4	US-09-540-236-3566
23	200	11.0	325	4	US-09-107-532A-4443
24	195.5	10.7	320	3	US-09-092-437-2
25	195.5	10.7	320	4	US-09-583-110-3858
26	195.5	10.7	320	4	US-09-107-433-3136
27	195	10.7	339	4	US-09-543-681A-5069

28	169.5	9.3	323	4	US-09-042-775-2	Sequence 2, Appli
29	167	9.2	343	3	US-09-134-001C-5484	Sequence 5484, Ap
30	163	9.0	334	4	US-09-134-000C-5726	Sequence 5726, Ap
31	161	8.8	311	4	US-09-710-279-1260	Sequence 1260, Ap
32	157	8.6	324	2	US-08-813-940-2	Sequence 2, Appli
33	156	8.6	391	4	US-09-489-039A-10121	Sequence 10121, A
34	153	8.4	104	4	US-09-230-485-5	Sequence 5, Appli
35	138	7.6	321	2	US-08-813-940-6	Sequence 6, Appli
36	127	7.0	52	3	US-08-917-299-21	Sequence 21, Appl
37	127	7.0	52	3	US-09-422-662-21	Sequence 21, Appl
38	127	7.0	52	4	US-09-730-763-21	Sequence 21, Appl
39	127	7.0	52	4	US-09-429-370-21	Sequence 21, Appl
40	103	5.7	555	3	US-09-173-151A-32	Sequence 32, Appl
41	100.5	5.5	382	4	US-09-134-000C-5889	Sequence 5889, Ap
42	99	5.4	233	3	US-08-931-825-2	Sequence 2, Appli
43	98.5	5.4	896	4	US-09-270-767-46130	Sequence 46130, A
44	95	5.2	233	3	US-09-161-662-2	Sequence 2, Appli
45	93.5	5.1	3830	4	US-09-693-205A-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-10-019-733-1
; Sequence 1, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 1
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-1

Query Match	100.0%;	Score 1821;	DB 4;	Length 351;
Best Local Similarity	100.0%;	Pred. No. 4e-189;		
Matches	351;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	MGDPERPEAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVFPQYDPDWKMYKQAO	60	
Db	1	MGDPERPEAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVFPQYDPDWKMYKQAO	60	
QY	61	ASFTAAEVDLSKDLPHWNKLKADKRYFISHILAFFAASDGI VNNELVERFSQEVQVPEA	120	
Db	61	ASFTAAEVDLSKDLPHWNKLKADKRYFISHILAFFAASDGI VNNELVERFSQEVQVPEA	120	
QY	121	RCFYGQIILNIENHSEMYSLIDTYINDPKREFLNAIETMPYVKKKADWALRIADRK	180	
Db	121	RCFYGQIILNIENHSEMYSLIDTYINDPKREFLNAIETMPYVKKKADWALRIADRK	180	
QY	181	STFGERVVAFAAVEGVFFSGSFAAI PWLKRGLMPGLTFSNELISRDEGLHCDFACLMFQ	240	
Db	181	STFGERVVAFAAVEGVFFSGSFAAI PWLKRGLMPGLTFSNELISRDEGLHCDFACLMFQ	240	
QY	241	YLVNKPSEERVREIIVDAVKIEQEFTEALPVGLIGNCILMKQYIEFVADRLVLVGLGFS	300	
Db	241	YLVNKPSEERVREIIVDAVKIEQEFTEALPVGLIGNCILMKQYIEFVADRLVLVGLGFS	300	
QY	301	KVFAENPDFMENISLEGKTNFFEKRVSEYQRPANVAETDNTDVTLDADF	351	

Db 301 KVFOAENPFDMENISLEGKTNFFKRVSEYQRFVMAETTDNVFTLDADF 351

RESULT 2

US-10-019-733-5
; Sequence 5, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 5
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-5

Query Match 79.8%; Score 1454; DB 4; Length 389;
Best Local Similarity 83.8%; Pred. No. 3.3e-149;
Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;

QY	31	EEPLLKSSRRFVIFPIQYDPDIWKYKQAOASFWTAEEVDLSKDLPHWNKLKADEKYPIS	90
Db	69	DEPLLRENPRFVIFPIEYHDIWQYKKAASFWTAEEVDLSKDIQHWESLKPEERYFIS	128
QY	91	HILAFFAASDGI NVNLFVSQVQVPEARCFYGFQIILNIHVHSEMYSLLLIDTYIRDPK	150
Db	129	HVLAFFAASDGI NVNLFVSQVQVPEARCFYGFQIILNIHVHSEMYSLLLIDTYIRDPK	188
QY	151	KREFLFNAIETMPYVKKADWALRWIADRKSTFGERVVAFVAAVEGFFSGSFAAIFWLKK	210
Db	189	EREFLFNAIETMPYVKKADWALRWIADRKSTFGERVVAFVAAVEGFFSGSFAAIFWLKK	248
QY	211	RGMLPGLTFSNELISRDEGLHCDPACLMFQYLVNKPSEERVREIIVDAVKIEQEFLEAL	270
Db	249	RGMLPGLTFSNELISRDEGLHCDPACLMFQYLVNKPSEERVREIIVDAVKIEQEFLEAL	308
QY	271	PVGLIGMNCILMKQYIEFVADRLVLGLFSGKVFOAENPFDMENISLEGKTNFFKRVSE	330
Db	309	PVKLIGMNCILMKQYIEFVADRLVLGLFSGKVFOAENPFDMENISLEGKTNFFKRVSE	368
QY	331	YQRFVMAETTDNVFTLDADF 351	
Db	369	YQRMGMSSPTENSFTLDADF 389	

RESULT 3

US-09-949-016-6016
; Sequence 6016, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6016
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6016

QY	31	EEPLLKSSRRFVIFPIQYDPDIWKYKQAOASFWTAEEVDLSKDLPHWNKLKADEKYPIS	90
Db	69	DEPLLRENPRFVIFPIEYHDIWQYKKAASFWTAEEVDLSKDIQHWESLKPEERYFIS	128
QY	91	HILAFFAASDGI NVNLFVSQVQVPEARCFYGFQIILNIHVHSEMYSLLLIDTYIRDPK	150
Db	129	HVLAFFAASDGI NVNLFVSQVQVPEARCFYGFQIILNIHVHSEMYSLLLIDTYIRDPK	188
QY	151	KREFLFNAIETMPYVKKADWALRWIADRKSTFGERVVAFVAAVEGFFSGSFAAIFWLKK	210
Db	189	EREFLFNAIETMPYVKKADWALRWIADRKSTFGERVVAFVAAVEGFFSGSFAAIFWLKK	248
QY	211	RGMLPGLTFSNELISRDEGLHCDPACLMFQYLVNKPSEERVREIIVDAVKIEQEFLEAL	270
Db	249	RGMLPGLTFSNELISRDEGLHCDPACLMFQYLVNKPSEERVREIIVDAVKIEQEFLEAL	308
QY	271	PVGLIGMNCILMKQYIEFVADRLVLGLFSGKVFOAENPFDMENISLEGKTNFFKRVSE	330
Db	309	PVKLIGMNCILMKQYIEFVADRLVLGLFSGKVFOAENPFDMENISLEGKTNFFKRVSE	368
QY	331	YQRFVMAETTDNVFTLDADF 351	
Db	369	YQRMGMSSPTENSFTLDADF 389	

RESULT 4

US-09-949-016-7896
; Sequence 7896, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7896
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7896

Query Match 79.8%; Score 1454; DB 4; Length 453;
Best Local Similarity 83.8%; Pred. No. 4.2e-149;
Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;

QY	31	EEPLLKSSRRFVIFPIQYDPDIWKYKQAOASFWTAEEVDLSKDLPHWNKLKADEKYPIS	90
Db	133	DEPLLRENPRFVIFPIEYHDIWQYKKAASFWTAEEVDLSKDIQHWESLKPEERYFIS	192
QY	91	HILAFFAASDGI NVNLFVSQVQVPEARCFYGFQIILNIHVHSEMYSLLLIDTYIRDPK	150

Db 193 HYLAFFAASDGVNENLVRFSEQVQITBARCFYQFOIAMENIHSEMYSLIIDTYIKDPK 252
QY 151 KREELFNAETMPYKKADWALRWADRSKSPGERVAFVAFVGVFFSGSPAIFWLKK 210
Db 253 REELFNAETMPYKKADWALRWADRSKSPGERVAFVAFVGVFFSGSPAIFWLKK 312
QY 211 RGLMPGLTFSNELISRDEGLHCDFACLMFQYLVNKPSEBVEREIIIVDAVKIEQEFLEAL 270
Db 313 RGLMPGLTFSNELISRDEGLHCDFACLMFQYLVNKPSEBVEREIIIVDAVKIEQEFLEAL 372
QY 271 PVGLIGMNCILMKQYIEFVADRLVELGFSKVQFAENPFDPMENISLEGKTNPFKRVSE 330
Db 373 PVKLGIMNCILMKQYIEFVADRLVELGFSKVQFAENPFDPMENISLEGKTNPFKRVSE 432
QY 331 YORFAVMAETTDNVTLDAF 351
Db 433 YORGMVMSPTENSFTLDAF 453

RESULT 5

US-09-248-796A-18017
; Sequence 18017, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 18017
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-18017

Query Match 62.1%; Score 1131; DB 4; Length 411;
Best Local Similarity 62.7%; Pred. No. 4.4e-114; Indels 6; Gaps 3;
Matches 210; Conservative 52; Mismatches 67

QY 23 NESEIKSNE--EPLLRKSSRRFVIPPYOYDPIWMYKQQAQSFVTAEEVDLSKDLPHW-N 79
Db 77 NRHLLKQESNEPLTTPDKTRHTIYPIKYPQLWQFYKSLASFVTAEEVDLSKDLPHW 136
QY 80 KKADEKYPISHILAFPAASDGVNENLVRFSEQVQVPEARCFYGFQILINHVSEMY 139
Db 137 KMNENEREFISRLVAFPAASDGVNENLVRFSEQVQVPEARCFYGFQILINHVSEMY 196
QY 140 LLDIVYIDPKKRELFNAETMPYKKADWALRWADRSKSPGERVAFVAFVGVFFSGSPA 199
Db 197 LIETIYFKDPEADFLNAIDNPIRKADWALRWADRSKSPGERVAFVAFVGVFFSGSPA 256
QY 200 GSFAAIFWKKRGLMPLGTLFNSNELISRDEGLHCDFACLMFQYLVNKPSEBVEREIIIVDA 259
Db 257 GSFAAIFWKKRGLMPLGTLFNSNELISRDEGLHCDFACLMFQYLVNKPSEBVEREIIIVDA 316
QY 260 KIEQFLEALPVGLIGMNCILMKQYIEFVADRLVELGFSKVQFAENPFDPMENISLEG 319
Db 317 DIEKEYFDALPVSLGIMNCILMKQYIEFVADRLVELGFSKVQFAENPFDPMENISLEG 376
QY 320 KTNPEKRVSEYQFAVMAET---TDNVFTLDAF 351
Db 377 KTNPEKRVSDYQKAGVMEKVENKNEKTGLPQDF 411

RESULT 6

US-08-307-499-19

; Sequence 19, Application US/08307499
; Patent No. 5651972
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; TITLE OF INVENTION: Live Vaccine Vector
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: U.S.A.
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/307,499
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,241
; FILING DATE: 1-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,630
; FILING DATE: 29-JUN-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/342,212
; FILING DATE: 21-APR-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF35.1.FWCC1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 274 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-307-499-19

Query Match 58.9%; Score 1072.5; DB 1; Length 274;
Best Local Similarity 75.6%; Pred. No. 5.4e-108;
Matches 204; Conservative 27; Mismatches 38; Indels 1; Gaps 1;
QY 32 EPLLKSSRRFVIPPYOYDPIWMYKQQAQSFVTAEEVDLSKDLPHWKNLKADEKYPISH 91
Db 2 EPILOESDSRFVIPPYKHYHDIMWKYKQSVASFVTAEEVDLSKDLPHWKNLKADEKYPISH 61
QY 92 ILAFAAASDGVNENLVRFSEQVQVPEARCFYGFQILINHVSEMYSLIIDTYIRDPK 151
Db 62 ILAFAAASDGVNENLVRFSEQVQVPEARCFYGFQILINHVSEMYSLIIDTYIRDPK 121
QY 152 REFLFNAETMPYKKADWALRWADRSKSPGERVAFVAFVGVFFSGSPAIFWLKKR 211
Db 122 KHLFNAETMECVKKKADWARKWISNK-VYGERVAFVAFVGVFFSGSPAIFWLKKR 180
QY 212 GLMPGLTFSNELISRDEGLHCDFACLMFQYLVNKPSEBVEREIIIVDAVKIEQEFLEALP 271
Db 181 GLMPGLTFSNELISRDEGLHCDFACLMFQYLVNKPSEBVEREIIIVDAVKIEQEFLEALP 240
QY 272 VGLIGMNCILMKQYIEFVADRLVELGFSK 301

Db 241 VDLIGNCCLMSQYIEFVADRLLTGCEK 270

RESULT 7

US-09-299-268-19
; Sequence 19, Application US/09299268
; Patent No. 6217882
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uella, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; TITLE OF INVENTION: Live Vaccine Vector
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: U.S.A.
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/299,268
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/901,127
; FILING DATE:
; APPLICATION NUMBER: US 07/908,241
; FILING DATE: 1-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,630
; FILING DATE: 29-JUN-1992
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/342,212
; FILING DATE: 21-APR-1992
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF35.1.FWCC1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 274 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-299-268-19

Query Match 58.9%; Score 1072.5; DB 3; Length 274;
Best Local Similarity 75.6%; Pred. No. 5.4e-108;
Matches 204; Conservative 27; Mismatches 38; Indels 1; Gaps 1;
QY 32 EPLLKSSRRFVFPFIQYDDIWKYKQAQASFTWAEVDLSKDLPHWNKLKADEKFIHSH 91
Db 2 EPILOESDRFVFPFIKYYDIWKYKQSVASFTWAEVDLSKDLDPDWKLTDEKFIKH 61
QY 92 ILAFAASDGIWENLVERFSQVPEARCFYGFQIILNENHSEMYSLIIDTYIRDPKK 151
Db 62 ILAFASSDGIWENLAEAFYVDQSEARCFYGFQIILNENHSEMYSLIIDTYVRDNIE 121
QY 152 REFLNATETMPYVKKADWALRWIADRKSTFGERVVAFAAEGVFFSGFAAIFWLKCR 211
Db 122 KQHLFNATETMECVKKADWARKWISSNK-VYGERVVAFAAEGVFFSGFAAIFWLKCR 180

QY 212 GLMPGLTFSNELISRDGHLHCDPACLMFOYLVNKPSEBRVREIIVDAVKIQEFLTEALP 271
Db 181 GLMPGLTFSNELISRDGHLHCDPACLMFKHLLHPPSPKEVITSIIIDAVNIKEFLTVAIP 240
QY 272 VGLIGNCILMKQYIEFVADRLLVGLGFSK 301
Db 241 VDLIGNCCLMSQYIEFVADRLLTGCEK 270

RESULT 8

US-08-136-743B-2
; Sequence 2, Application US/08136743B
; Patent No. 5459063
; GENERAL INFORMATION:
; APPLICANT: Barry S. Cooperman, Harvey Rubin,
; APPLICANT: Jerome Salem, and Alison L. Fisher
; TITLE OF INVENTION: "Plasmodium falciparum Ribonu-
; TITLE OF INVENTION: cleotide Reductase, DNA Sequences Therefor and Peptide inhibit
; TITLE OF INVENTION: Thereof"
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The University of Pennsylvania
; STREET: Suite 330
; STREET: 3700 Market Street
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19104-3246
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/136,743B
; FILING DATE: 10/14/93
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 3957-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: No. 5459063e
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 322 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-136-743B-2

Query Match 53.5%; Score 974.5; DB 1; Length 322;
Best Local Similarity 59.7%; Pred. No. 3e-97;
Matches 191; Conservative 51; Mismatches 75; Indels 3; Gaps 2;

QY 34 LIRKSSRRFVFPFIQYDDIWKYKQAQASFTWAEVDLSKDLPHWNKLKADEKFIHSHIL 93
Db 4 ILNKSDRFTLYPILYDPVFPFYKKAASFTWAEIDYSSDLKDFEKLNENEKFIKHVL 63
QY 94 AFAASDGIWENLVERFSQVPEARCFYGFQIILNENHSEMYSLIIDTYIRDPKRE 153
Db 64 AFAASDGIWENLAVSFLEVQITAEKPYFQIAVENHSEMYSLIIDTYIRDPKRE 123
QY 154 FLFNATETMPYVKKADWALRWIADRKSTFGERVVAFAAEGVFFSGFAAIFWLKRGKL 213
Db 124 NLFHALENTPAVKNAKALWAKWINDTNS-FAERIVANACVEGIFSSGFAIFWFKQNK 182
QY 214 MPGLTFSNELISRDGHLHCDPACLMFOYLVNKPSEBRVREIIVDA--VKIQEFLTEALP 271
Db 183 LHGLTFSNELISRDGHLHCDPACLMFOYLVNKPSEBRVREIIVDA--VKIQEFLTEALP 242

```
QY 272 VGLIGNCILMKQYIEFVADRLLVELGSKVFOAENPFDPMENISLEGKTNPFKEKRVSEY 331
Db 243 CDLIGNSRMLMSQTIEFVADRLLLECLGCKIPIHSKNPFWMDKISLQGTNPFKEKRVADY 302
QY 332 QRPVMAETTDNVFTLDADP 351
Db 303 OKSGVNAQRKDHVCLNTEF 322

RESULT 9
US-09-248-796A-18018
; Sequence 18018, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 18018
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-18018

Query Match 53.2%; Score 969.5; DB 4; Length 360;
Best Local Similarity 62.2%; Pred. No. 1.2e-96;
Matches 179; Conservative 53; Mismatches 53; Indels 3; Gaps 2;

QY 5 ERPEAAGLDQDERSSDTNESEBIK--SNBEPILKSSRRFVIPPPIQYPDIMKMYKQAQAS 62
Db 72 EKDDSLKHQDPLAKHVKHRLKQLEAEPLLVENKRRYVMPPIRYHINWFYKKAAS 131
QY 63 FWTAEVDLSKDLPHW-NKLKADEKYFISHILAFFAASDGI VNNENLVERFSOEYQVPEAR 121
Db 132 FWTAEIDLKDLDDNNKLNENERYFISRLVLAFFAASDGI VNGENLNFSTEVQLPEAK 191
QY 122 CYGFGQILLENVHSEMYSLIIDTYIRDPKRFELFNAIETMPYVKKADWALRWIADRS 181
Db 192 SYGFGQIMNTHSETYSLLIETIKDQPEADYLFNAIANIPCIQKADWAIKWIQDDEA 251
QY 182 TFGERVAPAAVEGVFFSGSFAAIFWLKKGRLMPLGTLFSGNELISRDEGLHCDPACLMFOY 241
Db 252 LYGERLVAPAAVEGVFFSGSFAIFWLKKGRLMPLGTLFSGNELICRDEGLHTDPACLLF 311
QY 242 LVNKPSEERREIIVDAVKIEQFLEALPVGLIGNCILMKQYIEFV 289
Db 312 LQNRPSPEIVERIITEAVDVEKYEYTDVLPVSLGLNCLMCPYIEFL 359

RESULT 10
US-08-905-223-358
; Sequence 358, Application US/08905223
; Patent No. 6222029
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste D.
; APPLICANT: Dueltel, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS
; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505

QY 272 VGLIGNCILMKQYIEFVADRLLVELGSKVFOAENPFDPMENISLEGKTNPFKEKRVSEY 331
Db 243 CDLIGNSRMLMSQTIEFVADRLLLECLGCKIPIHSKNPFWMDKISLQGTNPFKEKRVADY 302
QY 332 QRPVMAETTDNVFTLDADP 351
Db 303 OKSGVNAQRKDHVCLNTEF 322

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Win95
SOFTWARE: Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/905,223
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 358:
SEQUENCE CHARACTERISTICS:
LENGTH: 147 amino acids
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
ORIGINAL SOURCE:
ORGANISM: Homo Sapiens
TISSUE TYPE: Brain
FEATURE:
NAME/KEY: sig_peptide
LOCATION: -101...-1
IDENTIFICATION METHOD: Von Heijne matrix
OTHER INFORMATION: score 4.4
OTHER INFORMATION: seq ISHILAFFAASDG/IV
US-08-905-223-358

Query Match 42.2%; Score 768; DB 3; Length 147;
Best Local Similarity 100.0%; Pred. No. 2.5e-75;
Matches 147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIPPPIQYPDIMKMYKQAQ 60
Db 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIPPPIQYPDIMKMYKQAQ 60
QY 61 ASFTWAEVDLSKDLPHWNKLKADEKYFISHILAFFAASDGI VNNENLVERFSOEYQVPEA 120
Db 61 ASFTWAEVDLSKDLPHWNKLKADEKYFISHILAFFAASDGI VNNENLVERFSOEYQVPEA 120
QY 121 RCYFGQILLENVHSEMYSLIIDTYIR 147
Db 121 RCYFGQILLENVHSEMYSLIIDTYIR 147

RESULT 11
US-09-270-767-46236
; Sequence 46236, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46236
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-46236

Query Match 32.9%; Score 599; DB 4; Length 149;
Best Local Similarity 76.5%; Pred. No. 5.7e-57;
Matches 114; Conservative 14; Mismatches 21; Indels 0; Gaps 0;

QY 196 VFSGSFAAIFWLKKGRLMPLGTLFSGNELISRDEGLHCDPACLMFOYLVNKPSEERREI 255
```

Db 1 IFFSFSASIFWKKRGLMPLGTLFSNELISRDEGLHCDFAVLMFQHLVQRPKRRIIEII 60
QY 256 VDAVKIEOEFLTEALPVGLIGNCILMKOYIEFVADRLVLELGFVKVFOAENPPDFMEMI 315
Db 61 RDAVAIEOEFLTDALPVNLIGNCIDLMSQYIEFVADRLVLELGVGKIYNTKPNPFMEMI 120

QY 316 SLEGTNPFKEKVSQYQFAVNAETTDNV 344
Db 121 SLDGKTNPFKEKVGQYQVGMVSNPLDNV 149

RESULT 12

US-09-902-540-10006

; Sequence 10006, Application US/09902540

; Patent No. 6833447

; GENERAL INFORMATION:

; APPLICANT: Goldman, Barry S.

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Wiegand, Roger C.

; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

; FILE REFERENCE: 38-10(15849)B

; CURRENT APPLICATION NUMBER: US/09/902,540

; CURRENT FILING DATE: 2001-07-10

; PRIOR APPLICATION NUMBER: 60/217,883

; PRIOR FILING DATE: 2000-07-10

; NUMBER OF SEQ ID NOS: 16825

; SEQ ID NO 10006

; LENGTH: 323

; TYPE: PRT

; ORGANISM: Myxococcus xanthus

US-09-902-540-10006

Query Match

Best Local Similarity 27.2%; Score 495; DB 4; Length 323;

Matches 113; Conservative 63; Mismatches 125; Indels 22; Gaps 6;

QY 49 YPDIWKYKQAQASFWTAEEVDLSKDLPHW-NKLKADKCYFISHILAFPAASDGIENL 107
Db 3 YPQFEMRYNATKNTMTVEEDVFDSTDLVLRKMTDAERHLIHLVAFATGDSIVGNL 62

QY 108 VERFQEQVQPCRCYFGQIILNIENHSEMYSLIIDTYIRDPKREFLNATETMPYVKK 167
Db 63 VLNLKHLNAPARMYLSQLVEALHVFYLTLDYVDPDAERAMAAVNDNIPSIGR 122

QY 168 KADWALRWI-----ADRKSTFGER-----VVAFA-AVEGVFSGSFAAIFWLKRGGLMP 215
Db 123 KAQFCWKWMSIQGLDITLQTKAHRROFLNLCIFAGCIEGLFFFAFAVYVFLRSKGLN 182

QY 216 GLTFGNELISRDEGLHCDFACLMFQY-----LVNKPSEERVREIIVDAVKIEQEFUTE 268
Db 183 GLAAGTNVWFRDESAHMAFESIQVARKEEPLDFAQNERDVVAMLREAVECETOFAQD 242

QY 269 ALPVGLIGNCILMKOYIEFVADRLVLELGFVKVFOAENPPDFMENISLEGKTNPFKEV 328
Db 243 LLSGGVGLSVDQMRGYSYVADQRLQMLQVFGTKNPLHFMQLQDVQELTNFERRV 302

QY 329 SEYQRFVNAETTDNVFTLDAF 351
Db 303 SSYQVAVGGAATDVV--LDAAF 323

RESULT 13

US-09-328-352-4714

; Sequence 4714, Application US/09328352

; Patent No. 6562958

; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC99-03PA

; CURRENT APPLICATION NUMBER: US/09/328,352

; CURRENT FILING DATE: 1999-06-04

; NUMBER OF SEQ ID NOS: 8252

; SEQ ID NO 4714

; LENGTH: 445

; TYPE: PRT

; ORGANISM: Acinetobacter baumannii

US-09-328-352-4714

Query Match 21.0%; Score 381.5; DB 4; Length 445;

Best Local Similarity 27.8%; Pred. No. 1.2e-32;

Matches 97; Conservative 76; Mismatches 143; Indels 33; Gaps 10;

QY 9 AAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYPDWKYKQAQASFWTAEE 68
Db 98 APGLSELEMG-----QRQVDDKAMINCRADNLQLVPFKYEWAWQKYLDCGANNHMPQE 152

QY 69 VLSKDLPHW---NKLKADKCYFISHILAFPAASDGIENLVERFQEQVQPCRCYFG 125
Db 153 VNMNHDIALWKSENGLTEDERTIVMRSLGFFSTADSLVANNLVLAIYRHITNPECROYIL 212

QY 126 FOILNIENHSEMYSLIIDTYIRDPKREFLNATETMPYVKKADWALRW---IADRKST 182
Db 213 RQAFEEAIHTHAYQCYIESLGMDEGE---VFNMYREIPSVARKAAWGLKYTQSUSDPTFH 269

QY 183 FG-----ERVAFAAV-EGVFFSGSFAAIFWLKRGGLMPGLTFSNELISRDEGLHC 232
Db 270 TGTENDQRLRLNLIAYFCVLEGIFFYCGFTQLLSGMRNKGNGVAFQFYILRDESML 329

QY 233 DPACLMP-QYLVNKP-----BVRVRIIVDAVKIEQEFUTEALPVGLIGNCILMKOY 285
Db 330 NFGIDMINQIKIENPHLWTAEFQEQEVIQMLLEGTMLEIYARDTMPRGVLGMNMMBEY 389

QY 286 IEFVADRLVLELGFVKVFO-AENPPDFM-ENISLEGKTNPFKEVSEVQ 332
Db 390 LKFCNRRLSQLGPEQFAGVTNPFAMMSEMDLRKEKNFETRVTDYQ 438

RESULT 14

US-09-252-991A-19358

; Sequence 19358, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 19358

; LENGTH: 476

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-19358

Query Match 20.7%; Score 377.5; DB 4; Length 476;

Best Local Similarity 26.7%; Pred. No. 3.7e-32;

Matches 100; Conservative 72; Mismatches 154; Indels 49; Gaps 11;

QY 2 GDEPPEAAAGLDQDE---RSSDTNESEIKSNEEPLLRKSSRRFV-----I 44
Db 99 GSVEEARAVSADSDAVARAKKALNDLDIQEGLDLDSGAARVQVQDKQMINARADLNQL 158

QY 45 FPIQYPDWKYKQAQASFWTAEEVDLSKDLPHWNK---LKADEKYFISHILAFPAASDG 101
Db 159 VPKYDWAWQKYLDDGCGANNHMPQEVNMNADIALWKSDGLSEHERRIVMRNLGFFSTADS 218

QY 102 IVNENLVERFSQEQVQPCRCYFGQIILNIENHSEMYSLIIDTYIRDPKREFLNATET 161

```
Db 219 LVANNLVAVYRLITNPECRQVILRQAFEEAHTHAYQYCIESLGMDEGE---IFNNYHE 275
QY 162 MPYVKKKADWALRWI-----ADRKSTGERVVAFAV-EGVFFSGSFAAIF 206
Db 276 IPSVAKKASWGLKYFESISDPMPQTGTPETDRQ--FLNLIAYYCVLEGIFFCGFTQIL 333
QY 207 WLKKGGLMPLGFTSNEISRLSDGLHCDPAC-LMFQYLVNKP-----SBERVREIIVDAV 259
Db 334 SMGRNKMGTGAEQFYILRDESMHNFIDVINQIKIENPHLWDAQMKDEATQMLQGT 393
QY 260 KTEQFLTEALPVGLIGMNCILMKQYIEFVADRLVELGFSKVQ-AENPPDFMENI-SL 317
Db 394 QLEIEYARDTMRGVLGNANMMDYLKFIANRLTQIGLKEEYPGTTPFPWMSBIMDL 453
QY 318 EGTKNFFEKRVSEYQ 332
Db 454 KKEKNFFETRVIEYQ 468

RESULT 15
US-09-198-452A-1059
; Sequence 1059 Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffaiss, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 1059
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-1059

Query Match 18.4%; Score 335; DB 4; Length 365;
Best Local Similarity 26.6%; Pred. No. 1e-27;
Matches 84; Conservative 64; Mismatches 134; Indels 34; Gaps 6;

QY 44 IPFIQYDIIKMYKQAQAFWTAEEVDLSKDLPHW--NKLKADKDYFISHILAFPAASDG 101
Db 50 LVPIKYNWHEHYLNGCANNWLPTVPWARDLWKSDELSEDERKVVILLNLGFFSTAES 109
QY 102 IVNENLVERFSQEVQPEARCFYQIILTIENVHSEMYSLIDTIYRDPKRBFLFNAIET 161
Db 110 LVGNIVLAIFFGHITNPEARQYLLRQAFEEAVHTHTFLYICESLGLDEGE---VFNAYNE 166
QY 162 MPYVKKKADWALRWIAD-----RKSTPGERVVAFAV-EGVFFSGSFAA 204
Db 167 RASIRAKDDFQMTLTVDLDPNFSQSSBGLQGFKNLVGVYII---MEGIFFYSGFWM 222
QY 205 IPWLKRGMLPGLTFSNELISDEGLHCDPACLMFQYLVNKPSE-----ERVREIIVD 257
Db 223 ILSFHRQNMGTGIGQYQYILRDETHLNFIDLINGIKEENPEVMTTELQEEIVALTEK 282
QY 258 AVKIQEFLTEALPVGLIGMNCILMKQYIEFVADRLVELGFSKVQAFENPFDFM-ENIS 316
Db 283 AVELEIEYAKDCLPGILGLRSSMFDIVRHIADRLRLRIGLKPIYHSRNPFPWMSBIMDL 342
QY 317 LEGKTNFFEKRVSEYQ 332
Db 343 LNKKNFFETRVIEYQ 358
```

Search completed: October 30, 2005, 06:58:42
Job time : 25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2005, 06:57:41 ; Search time 72 Seconds
(without alignments)
2037.960 Million cell updates/sec

Title: US-10-698-228-1
Perfect score: 1821
Sequence: 1 MGDPERPEAAGLDQDERSSS.....QRFVMAETDNDVFTLDADF 351

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1865214 seqs, 418043040 residues

Total number of hits satisfying chosen parameters: 1865214

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1821	100.0	351	15	US-10-698-228-1
2	1454	79.8	389	15	US-10-698-228-5
3	1454	79.8	389	16	US-10-408-765A-559
4	1454	79.8	389	16	US-10-733-878-457
5	1454	79.8	389	18	US-10-756-149-4899
6	1454	79.8	413	9	US-09-925-301-1347
7	1454	79.8	453	18	US-10-220-335-686
8	1454	79.8	453	18	US-10-450-763-45646
9	1425.5	78.3	386	15	US-10-403-571-76
10	1348	74.0	393	20	US-11-097-143-19182
11	1242	68.2	430	15	US-10-369-493-4122

12	1230	67.5	405	14	US-10-128-714-8245	Sequence 8245, Ap
13	1214	66.7	381	15	US-10-369-493-5686	Sequence 5686, Ap
14	1192.5	65.5	381	14	US-10-128-714-3245	Sequence 3245, Ap
15	1192	65.5	413	14	US-10-032-585-7537	Sequence 7537, Ap
16	1176.5	64.6	321	15	US-10-369-493-10365	Sequence 10365, A
17	1168.5	64.2	399	15	US-10-369-493-22170	Sequence 22170, A
18	1161.5	63.8	391	15	US-10-369-493-22565	Sequence 22565, A
19	1161.5	63.8	401	15	US-10-320-797-3145	Sequence 3145, Ap
20	1147.5	63.0	329	15	US-10-424-599-215866	Sequence 215866, A
21	1133.5	62.2	717	16	US-10-425-115-215845	Sequence 215845, A
22	1132.5	62.2	345	16	US-10-437-963-113592	Sequence 113592, A
23	1128.5	62.0	339	16	US-10-767-701-45287	Sequence 45287, A
24	1125.5	61.8	339	16	US-10-425-115-247984	Sequence 247984, A
25	1114.5	61.2	329	15	US-10-424-599-269685	Sequence 269685, A
26	1097.5	60.3	339	15	US-10-424-599-183313	Sequence 183313, A
27	1092.5	60.0	400	16	US-10-425-115-200125	Sequence 200125, A
28	1061	58.3	347	16	US-10-739-930-6922	Sequence 6922, Ap
29	916	50.3	270	18	US-10-220-335-342	Sequence 342, App
30	871.5	47.9	344	15	US-10-369-493-13401	Sequence 13401, A
31	840	46.1	254	16	US-10-425-115-215843	Sequence 215843, A
32	805	44.2	345	15	US-10-369-493-21987	Sequence 21987, A
33	784	43.1	225	16	US-10-425-115-215840	Sequence 215840, A
34	701	38.5	138	15	US-10-264-237-2195	Sequence 2195, Ap
35	680	37.3	335	18	US-10-450-763-45645	Sequence 45645, A
36	662	36.4	181	18	US-10-450-763-37219	Sequence 37219, A
37	564.5	31.0	338	14	US-10-156-761-10564	Sequence 10564, A
38	544	29.9	156	15	US-10-424-599-215870	Sequence 215870, A
39	491	27.0	307	15	US-10-369-493-19363	Sequence 19363, A
40	456.5	25.1	145	16	US-10-767-701-45286	Sequence 45286, A
41	397	21.8	398	15	US-10-282-122A-51357	Sequence 51357, A
42	384.5	21.1	415	15	US-10-282-122A-66252	Sequence 66252, A
43	384	21.1	403	15	US-10-282-122A-50118	Sequence 50118, A
44	383.5	21.1	415	15	US-10-282-122A-69814	Sequence 69814, A
45	380.5	20.9	416	15	US-10-282-122A-68043	Sequence 68043, A

ALIGNMENTS

RESULT 1

US-10-698-228-1

Sequence 1, Application US/10698228

Publication No. US2004007253A1

GENERAL INFORMATION:

APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA

TITLE OF INVENTION: New Protein and its DNA

FILE REFERENCE: 2619WOOP

CURRENT APPLICATION NUMBER: US/10/698,228

CURRENT FILING DATE: 2003-10-30

PRIOR APPLICATION NUMBER: US/10/019,733

PRIOR FILING DATE: 2001-12-28

PRIOR APPLICATION NUMBER: JP 11-181131

PRIOR FILING DATE: 1999-06-28

PRIOR APPLICATION NUMBER: JP 11-192391

PRIOR FILING DATE: 1999-07-06

PRIOR APPLICATION NUMBER: JP 2000-017770

PRIOR FILING DATE: 2000-01-21

NUMBER OF SEQ ID NOS: 14

SEQ ID NO 1

LENGTH: 351

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION:

US-10-698-228-1

Query Match 100.0%; Score 1821; DB 15; Length 351;

Best Local Similarity 100.0%; Pred. No. 1.7e-162;

Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVFPIQYDPDWKMYKQAO 60

DB 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVFPIQYDPDWKMYKQAO 60

QY 61 ASFWTAEEVDLSKDLPHWNKLKADKDYFISHILAFPAASDGIIVNENLVERFSQEVQVPEA 120
DB 61 ASFWTAEEVDLSKDLPHWNKLKADKDYFISHILAFPAASDGIIVNENLVERFSQEVQVPEA 120
QY 121 RCFYGFQILNIENHSMYSLIDTYIRDPKREFFLNAIETMPYVKKKADWALRWIADRK 180
DB 121 RCFYGFQILNIENHSMYSLIDTYIRDPKREFFLNAIETMPYVKKKADWALRWIADRK 180
QY 181 STFGERVVAFVAVGVFFSGSFAAIFWLKRGMLPGLTFSNELISRDGLHCDPACLMFQ 240
DB 181 STFGERVVAFVAVGVFFSGSFAAIFWLKRGMLPGLTFSNELISRDGLHCDPACLMFQ 240
QY 241 YLVNKPSEBVRREIIVDAVKIEOEFTALPVGLIGMNCILMKQYIEFVADRLVLVGF 300
DB 241 YLVNKPSEBVRREIIVDAVKIEOEFTALPVGLIGMNCILMKQYIEFVADRLVLVGF 300
QY 301 KVFOAENPDDFMENISLEGKTNPFKRVSEYQRFVMAETTDNVFTLDADF 351
DB 301 KVFOAENPDDFMENISLEGKTNPFKRVSEYQRFVMAETTDNVFTLDADF 351

RESULT 2

US-10-698-228-5
; Sequence 5, Application US/10698228
; Publication No. US20040072253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; PRIOR FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 5
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-5

Query Match 79.8%; Score 1454; DB 15; Length 389;
Best Local Similarity 83.8%; Pred. No. 6.6e-128;
Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;

QY 31 EBPRLKSSRRFVIFPIQYDWMKYKQASFWTAEEVDLSKDLPHWNKLKADKDYFIS 90
DB 69 DEPLLENPRRFVIFPIEYHDIMQYKKAESFWTAEEVDLSKDIQHWSLKPEERYFIS 128
QY 91 HILAPFAASDGIIVNENLVERFSQEVQVPEARCFYGFQILNIENHSMYSLIDTYIRDPK 150
DB 129 HVLAPFAASDGIIVNENLVERFSQEVQVTEARCFYGFQIAMENIHSEMYSLLIDTYIKDPK 188
QY 151 KREFLNAIETMPYVKKKADWALRWIADRKSTGERVVAFAAVEGVFFSGSFAAIFWLKK 210
DB 189 EREFLNAIETMPYVKKKADWALRWIADRKSTGERVVAFAAVEGVFFSGSFAAIFWLKK 248
QY 211 RGLMPGLTFSNELISRDGLHCDPACLMFQYLVNKPSEBVRREIIVDAVKIEOEFTAL 270
DB 249 RGLMPGLTFSNELISRDGLHCDPACLMFQYLVNKPSEBVRREIIVDAVKIEOEFTAL 308
QY 271 PVGLIGMNCILMKQYIEFVADRLVLVGFSGKVFQAEENPDDFMENISLEGKTNPFKRVSE 330
DB 309 PVKLGIMNCTLMKQYIEFVADRLVLVGFSGKVFQAEENPDDFMENISLEGKTNPFKRVSE 368

QY 331 YQFVAVMAETTDNVFTLDADF 351
DB 369 YQMGVMSSTPENSFTLDADF 389

RESULT 3

US-10-408-765A-559
; Sequence 559, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Watson, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-559

Query Match 79.8%; Score 1454; DB 16; Length 389;
Best Local Similarity 83.8%; Pred. No. 6.6e-128;
Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;

QY 31 EBPRLKSSRRFVIFPIQYDWMKYKQASFWTAEEVDLSKDLPHWNKLKADKDYFIS 90
DB 69 DEPLLENPRRFVIFPIEYHDIMQYKKAESFWTAEEVDLSKDIQHWSLKPEERYFIS 128
QY 91 HILAPFAASDGIIVNENLVERFSQEVQVPEARCFYGFQILNIENHSMYSLIDTYIRDPK 150
DB 129 HVLAPFAASDGIIVNENLVERFSQEVQVTEARCFYGFQIAMENIHSEMYSLLIDTYIKDPK 188
QY 151 KREFLNAIETMPYVKKKADWALRWIADRKSTGERVVAFAAVEGVFFSGSFAAIFWLKK 210
DB 189 EREFLNAIETMPYVKKKADWALRWIADRKSTGERVVAFAAVEGVFFSGSFAAIFWLKK 248
QY 211 RGLMPGLTFSNELISRDGLHCDPACLMFQYLVNKPSEBVRREIIVDAVKIEOEFTAL 270
DB 249 RGLMPGLTFSNELISRDGLHCDPACLMFQYLVNKPSEBVRREIIVDAVKIEOEFTAL 308
QY 271 PVGLIGMNCILMKQYIEFVADRLVLVGFSGKVFQAEENPDDFMENISLEGKTNPFKRVSE 330
DB 309 PVKLGIMNCTLMKQYIEFVADRLVLVGFSGKVFQAEENPDDFMENISLEGKTNPFKRVSE 368
QY 331 YQFVAVMAETTDNVFTLDADF 351
DB 369 YQMGVMSSTPENSFTLDADF 389

RESULT 4

US-10-733-878-457
; Sequence 457, Application US/10733878
; Publication No. US20040224408A1
; GENERAL INFORMATION:
; APPLICANT: Jean-Philippe Girard
; APPLICANT: Francois Amalric
; APPLICANT: Myriam Roussigne
; APPLICANT: Thomas Clouaire
; TITLE OF INVENTION: THAP PROTEINS AS NUCLEAR RECEPTORS FOR
; TITLE OF INVENTION: CHEMOKINES AND ROLES IN TRANSCRIPTIONAL REGULATION, CELL
; FILE REFERENCE: BIOBANK.012A
; CURRENT APPLICATION NUMBER: US/10/733,878

```
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 60/432699
; PRIOR FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/485027
; PRIOR FILING DATE: 2003-07-03
; NUMBER OF SEQ ID NOS: 535
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 457
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-733-878-457

Query Match
Best Local Similarity 79.8%; Score 1454; DB 16; Length 389;
Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;

QY 31 BEPLLRKSRFRVIFPIQYPDIMWKYKQAQSFWTAEVDLSKDLPHNKLKADKEYFIS 90
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
69 DEPLLRENPRFRVIFPIEYHDIMQYKKAESFWTAEVDLSKDIQHWESLKPPEERYFIS 128
QY 91 HILAFPAASDGIENLVRFSGQVQVPEARCFYGFQIILNVHSEMYSLIDITYIRDPK 150
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
129 HVLAFPAASDGIENLVRFSGQVQVTEARCFYGFQIAMIHSEMYSLIDITYIKDPK 188
QY 151 KREFLFNAIETMPYVKKKADWALRWIADRKSTFGERVVAPAAVEGVPFSGSFAAIFWLKK 210
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
189 EREFLFNAIETMPCVKKKADWALRWIGDKKATYGERVVAFAAVEGIFPFGSFAIFWLKK 248
QY 211 RGLMPGLTFSNELISRDEGLHCDFACLMFQYLVNKPSEERVREIIVDAVKIEQEFTEAL 270
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
309 PVKLGIMNCTLMKQYIEFVADRLMLGLGFSKVFVRNPFDFMENISLEGKTNFFPKRVSE 330
QY 331 YORFVMAETTDNVFTLDADF 351
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
369 YORGMVNSPTENSFTLDADF 389

Query Match
Best Local Similarity 79.8%; Score 1454; DB 9; Length 413;
Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;

QY 31 BEPLLRKSRFRVIFPIQYPDIMWKYKQAQSFWTAEVDLSKDLPHNKLKADKEYFIS 90
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
93 DEPLLRENPRFRVIFPIEYHDIMQYKKAESFWTAEVDLSKDIQHWESLKPPEERYFIS 152
QY 91 HILAFPAASDGIENLVRFSGQVQVPEARCFYGFQIILNVHSEMYSLIDITYIRDPK 150
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
153 HVLAFPAASDGIENLVRFSGQVQVTEARCFYGFQIAMIHSEMYSLIDITYIKDPK 212
QY 211 KREFLFNAIETMPYVKKKADWALRWIADRKSTFGERVVAPAAVEGVPFSGSFAAIFWLKK 210
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
213 EREFLFNAIETMPCVKKKADWALRWIGDKKATYGERVVAFAAVEGIFPFGSFAIFWLKK 272
QY 211 RGLMPGLTFSNELISRDEGLHCDFACLMFQYLVNKPSEERVREIIVDAVKIEQEFTEAL 270
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
273 RGLMPGLTFSNELISRDEGLHCDFACLMFQYLVNKPSEERVREIIVDAVKIEQEFTEAL 332
QY 271 PVKLGIMNCTLMKQYIEFVADRLMLGLGFSKVFVRNPFDFMENISLEGKTNFFPKRVSE 330
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
333 PVKLGIMNCTLMKQYIEFVADRLMLGLGFSKVFVRNPFDFMENISLEGKTNFFPKRVSE 392
QY 331 YORFVMAETTDNVFTLDADF 351
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
393 YORGMVNSPTENSFTLDADF 413

RESULT 7
US-10-220-335-686
; Sequence 686, Application US/10220335
; Publication No. US20050175607A1
```

```
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 60/432699
; PRIOR FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/485027
; PRIOR FILING DATE: 2003-07-03
; NUMBER OF SEQ ID NOS: 535
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 457
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-733-878-457

Query Match
Best Local Similarity 79.8%; Score 1454; DB 16; Length 389;
Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;

QY 31 BEPLLRKSRFRVIFPIQYPDIMWKYKQAQSFWTAEVDLSKDLPHNKLKADKEYFIS 90
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
69 DEPLLRENPRFRVIFPIEYHDIMQYKKAESFWTAEVDLSKDIQHWESLKPPEERYFIS 128
QY 91 HILAFPAASDGIENLVRFSGQVQVPEARCFYGFQIILNVHSEMYSLIDITYIRDPK 150
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
129 HVLAFPAASDGIENLVRFSGQVQVTEARCFYGFQIAMIHSEMYSLIDITYIKDPK 188
QY 151 KREFLFNAIETMPYVKKKADWALRWIADRKSTFGERVVAPAAVEGVPFSGSFAAIFWLKK 210
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
189 EREFLFNAIETMPCVKKKADWALRWIGDKKATYGERVVAFAAVEGIFPFGSFAIFWLKK 248
QY 211 RGLMPGLTFSNELISRDEGLHCDFACLMFQYLVNKPSEERVREIIVDAVKIEQEFTEAL 270
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
249 RGLMPGLTFSNELISRDEGLHCDFACLMFQYLVNKPSEERVREIIVDAVKIEQEFTEAL 308
QY 271 PVKLGIMNCTLMKQYIEFVADRLMLGLGFSKVFVRNPFDFMENISLEGKTNFFPKRVSE 330
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
309 PVKLGIMNCTLMKQYIEFVADRLMLGLGFSKVFVRNPFDFMENISLEGKTNFFPKRVSE 368
QY 331 YORFVMAETTDNVFTLDADF 351
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
369 YORGMVNSPTENSFTLDADF 389

RESULT 5
US-10-756-149-4899
; Sequence 4899, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/756,149
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4899
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-756-149-4899

Query Match
Best Local Similarity 83.8%; Score 1454; DB 18; Length 389;
Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;

QY 31 BEPLLRKSRFRVIFPIQYPDIMWKYKQAQSFWTAEVDLSKDLPHNKLKADKEYFIS 90
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
69 DEPLLRENPRFRVIFPIEYHDIMQYKKAESFWTAEVDLSKDIQHWESLKPPEERYFIS 128
QY 91 HILAFPAASDGIENLVRFSGQVQVPEARCFYGFQIILNVHSEMYSLIDITYIRDPK 150
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
129 HVLAFPAASDGIENLVRFSGQVQVTEARCFYGFQIAMIHSEMYSLIDITYIKDPK 188
```


Db 47 RKIPDESGQKAKGAVEEPLLENKPHRFVIFPIQYHDIWQYKKAASFTWAEVDL 106
Qy 72 SKDLPHWKLKADKDYFISHILAFPAASDGIWNLVERFSEVQVPEARCFVGFQIILE 131
Db 107 SKDLQWLSLKDDEERYFISHVLAFAASDGIWNLVERFSEVQVTEARCFVGFQIAME 166
Qy 132 NVHSEMYSLLDITYIRDPKREFLFNAIETMPYVKKKADWALRWIADRKSTFGERVVAPA 191
Db 167 NIHSEMYSLLDITYIKDSKEREFLFNAIETMPCVKKADWALNWIQGNARYGERVVAFA 226
Qy 192 AVEGVFFSGSPAIAFWLKKRGLMPGLTFSNELISRDEGLHCDPACLMFOYLVNKPSEERV 251
Db 227 AVEGVFFSGSPASIFWLKKRGLMPGLTFSNELISRDEGLHCDPACLMFKHLINKPSEETV 286
Qy 252 REIIVDAVKIEOEFTALPVLGIGNCILMKOYIEFVADRLVLELGFSEVQVFAENPDPF 311
Db 287 KKIINVAIEOEFTDALPVKLIGNCDLMKQYIEFVADRLVLELGFSEVQVFAENPDPF 346
Qy 312 MENISLEGKTNFFKRVSEYQRFVMAETTDNVFTILDADP 351
Db 347 MENISLEGKTNFFKRVSEYQRFVMAETTDNVFTILDADP 386

RESULT 10

US-11-097-143-19182
; Sequence 19182, Application US/11097143
; Publication No. US2005020858A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19182
; LENGTH: 393
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-19182

Query Match 74.0%; Score 1348; DB 20; Length 393;
Best Local Similarity 73.9%; Pred. No. 6.3e-118;
Matches 252; Conservative 41; Mismatches 48; Indels 0; Gaps 0;
Qy 11 GLDQDERSSTNESEIKSNEBPLRKSSRRFVIFPIQYDPDKWYKQAASFTWAEVD 70
Db 53 GIGKANSIMKSVTFPPDSLEPLRENPFRVIFPIQYHDIWQYKKAASFTWAEVD 112
Qy 71 LSKDLPHWKLKADKDYFISHILAFPAASDGIWNLVERFSEVQVPEARCFVGFQIILI 130
Db 113 LSKDLTDWHLKDDERHSHVLAFAASDGIWNLVERFSEVQVTEARCFVGFQIAM 172
Qy 131 ENVHSEMYSLLDITYIRDPKREFLFNAIETMPYVKKKADWALRWIADRKSTFGERVVAP 190

Db 173 ENVHSEMYSLLDITYIRDPKREFLFNAIETMPYVKKKADWALRWIADRKSTFGERVVAP 232
Qy 191 AAVEGVFFSGSPAIAFWLKKRGLMPGLTFSNELISRDEGLHCDPACLMFOYLVNKPSEER 250
Db 233 AAVEGVFFSGSPASIFWLKKRGLMPGLTFSNELISRDEGLHCDPACLMFOYLVNKPSEER 292
Qy 251 VREIIVDAVKIEOEFTALPVLGIGNCILMKOYIEFVADRLVLELGFSEVQVFAENPDPF 310
Db 293 IIEIIVDAVAIEOEFTDALPVKLIGNCDLMKQYIEFVADRLVLELGFSEVQVFAENPDPF 352
Qy 311 FMENISLEGKTNFFKRVSEYQRFVMAETTDNVFTILDADP 351
Db 353 FMENISLEGKTNFFKRVSEYQRFVMAETTDNVFTILDADP 393

RESULT 11

US-10-369-493-4122
; Sequence 4122, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 4122
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Neurospora crassa
US-10-369-493-4122

Query Match 68.2%; Score 1242; DB 15; Length 430;
Best Local Similarity 62.6%; Pred. No. 6.7e-108;
Matches 243; Conservative 52; Mismatches 47; Indels 46; Gaps 6;
Qy 10 AGLDQDERSSTNESEIKSNEBPLRKSSRRFVIFPIQYDPDKWYKQAASFTWAEVD 51
Db 43 AEIDAENANKKAAEAKQWAPTLKPEANEPLTENPQRFVLPFIKYHEVCHASQTFTT 102
Qy 52 IWKYKQAASFTWAEVDLSKDLPHW-NKLKADKDYFISHILAPF 96
Db 103 RKVVHGLIGYPHYRIWQYKKAASFTWAEIDLKDLHWNRLNDEKFFISHILAPF 162
Qy 97 AASDGIWNLVERFSEVQVPEARCFVGFQIILENVHSEMYSLLDITYIRDPKREFLF 156
Db 163 AASDGIWNLVERFSEVQVPEARCFVGFQIILENVHSEMYSLLDITYIRDPKREFLF 222
Qy 157 NAIETMPYVKKKADWALRWIADRKSTFGERVVAPAAVEGVFFSGSPAIAFWLKKRGLMPG 216
Db 223 NAIDTIPICRKADWALRWITDKSSTFAQRILVAPAAVEGVFFSGSPAIAFWLKKRGLMPG 282
Qy 217 LTFSENLISRDEGLHCDPACLMFOYLVNKPSEERVRIIVDAVKIEOEFTALPVLGIL 276
Db 283 LTFSENLISRDEGLHCDPACLMFOYLVNKPSEERVRIIVDAVKIEOEFTALPVLGIL 342
Qy 277 MNCILMKOYIEFVADRLVLELGFSEVQVFAENPDPFMENISLEGKTNFFKRVSEYQVFA 336
Db 343 MNAIDLKQYIEFVADRLVLELGFSEVQVFAENPDPFMENISLEGKTNFFKRVSEYQVFA 402
Qy 337 M-----AETTDNV-----FTILDADP 351
Db 403 MNSTKKADADA EVAKNENGSGDFTFDEDF 430

RESULT 12

US-10-128-714-8245
; Sequence 8245, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wengqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in *Aspergillus fumigatus* and
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; PRIOR FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8245
; LENGTH: 405
; TYPE: PRT
; ORGANISM: *Aspergillus fumigatus*
US-10-128-714-8245

Query Match 67.5%; Score 1230; DB 14; Length 405;

Best Local Similarity 67.0%; Pred. No. 8.2e-107;
Matches 234; Conservative 52; Mismatches 49; Indels 14; Gaps 3;

QY 16 EBSDDTNESEIKSNEEPLLRKSSRRFVFPPIQYDIPWIKMYKQAQASFWTAEVDLSKDL 75
DB 58 EASKVAQSKELEAN-EPILQENPHRFVLPPIKYHEIWQMYKAEASFWTAEIDLSKDL 116

QY 76 PHW-NKLKADKDYFISHILAFPAASDGIYVNEVLVERFQVPEARCFYGFQILLNVH 134
DB 117 HDNNELNDDERYFISHVLAFAASDGIYVNEVLVERFQVPEARCFYGFQIMLENH 176

QY 135 SEMYSLIDITYIRDPKKRBFLENAIETMPYVKKADWALRWIADRKSTTGERVVAFAVE 194
DB 177 SETYSLIDITYIKEPKORTYLFDAIDVPCIRKKAQWAMRWISDKBSTFAQLVAFAAVE 236

QY 195 GYFSGSFAAIFWLKKGMLPGLTFSNELISRDEGLHCDFACLMFOYLVNKPSEERVREI 254
DB 237 GIFFSGSFASIFWLKKGMLPGLTFSNELISRDEGLHCDFACLMFOYLVNKPSEERVREI 296

QY 255 IVDVAKIEQEFLEALPVGLIGNCTILMKQYIEFVADRLLLVELGFSKVFQAEPPDFMEN 314
DB 297 IVEVAIEQEFLEALPVGLIGNCTILMKQYIEFVADRLLLVELGFSKVFQAEPPDFMES 356

QY 315 ISLEGKTNFPEKRVSEYQRFVMAET-----TDNVFTLDADF 351
DB 357 ISLAGKTNFPEKRVSEYQRFVMAETKQASQDTAKDANGGLCFDEDF 405

RESULT 13

US-10-369-493-5686
; Sequence 5686, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkie, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 5686
; LENGTH: 381
; TYPE: PRT
; ORGANISM: *Caenorhabditis elegans*
US-10-369-493-5686

Query Match 66.7%; Score 1214; DB 15; Length 381;

Best Local Similarity 67.2%; Pred. No. 2.4e-105;
Matches 231; Conservative 48; Mismatches 61; Indels 4; Gaps 2;

QY 12 LDQERSSD--TNESEIK--SNEEPLLRKSSRRFVFPPIQYDIPWIKMYKQAQASFWTAE 67
DB 38 VDQTKAASAEETNNESEVNELDADPEMLQDLNRFVIFPLKHHDINWYKXAVASFWTVE 97

QY 68 EVDLSKDLPHWNKLKADKDYFISHILAFPAASDGIYVNEVLVERFQVPEARCFYGFQ 127
DB 98 EVDLKGDMNDWEXMGDEQYFISILAFPAASDGIYVNEVLVERFQVPEARCFYGFQ 157

QY 128 ILIENVHSEMYSLIDITYIRDPKKRBFLENAIETMPYVKKADWALRWIADRKSTGERV 187
DB 158 IAIENHSEMYSKLIETIIRDETERTLNAVDEFFIKKADWALRWISDKKASFAERL 217

QY 188 VAPAAVEGVFFSGSFAAIFWLKKGMLPGLTFSNELISRDEGLHCDFACLMFOYLVNKP 247
DB 218 IAPAAVEGVFFSGSFAAIFWLKKGMLPGLTFSNELISRDEGLHCDFACLMFOYLVNKP 277

QY 248 EERVREIIVDAVIEQEFLEALPVGLIGNCTILMKQYIEFVADRLLLVELGFSKVFQAE 307
DB 278 QQRIYDIIDAVAIEQEFLEALPVDMIGNCRLMSQYIEFVADRLLLVELGCDLYKSKN 337

QY 308 PPDFMENISLEGKTNFPEKRVSEYQRFVMAETTDNVFTLDADF 351
DB 338 PPDFMENISIDGKTNFPEKRVSEYQRFVGMVMEARQFDLEADF 381

RESULT 14

US-10-128-714-3245
; Sequence 3245, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wengqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in *Aspergillus fumigatus* and
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; PRIOR FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3245
; LENGTH: 381

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 30, 2005, 06:58:47 ; Search time 3469 Seconds

(without alignments)
4902.791 Million cell updates/sec

Title: US-10-698-228-1

Perfect score: 1821

Sequence: 1 MGDPERPEAAGLDQDRSSS.....QRPVMAETTDNVFTLDADF 351

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10698228/runat_26102005_100714_5107/app.query.fasta.1.519
-DB=GenEmbl -QFMT=fastcap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10698228@cgn_1_1_2527@runat_26102005_100714_5107 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOB=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb_ba.*
2: gb_hg.*
3: gb_in.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1821	100.0	1053	6	AR454868
2	1821	100.0	1053	6	AR454877
3	1821	100.0	1053	6	BD064764
4	1821	100.0	1053	6	BD064773

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

5	1821	100.0	1053	6	BD093077
6	1821	100.0	1053	6	BD093086
7	1821	100.0	1056	6	CQ714252
8	1821	100.0	1056	6	AB166671
9	1821	100.0	1081	6	AR454870
10	1821	100.0	1081	6	BD064766
11	1821	100.0	1081	6	BD093079
12	1821	100.0	4955	6	AR454869
13	1821	100.0	4955	6	BD064765
14	1821	100.0	4955	6	BD093078
15	1821	100.0	4955	9	AB036063
16	1817	99.8	1601	6	BD156916
17	1817	99.8	1601	6	AX877905
18	1817	99.8	1601	9	AK001965
19	1725	94.7	4532	10	BC058103
20	1627.5	89.4	171737	10	AC122379
21	1514	83.1	900	9	AB163437
22	1474	80.9	4571	9	BC042468
23	1460.5	80.2	1649	9	BC001886
24	1460.5	80.2	1653	9	BC030154
25	1460.5	80.2	1794	9	S40301
26	1460.5	80.2	2500	6	AR300103
27	1460.5	80.2	2500	6	AR380825
28	1460.5	80.2	2500	6	AR438503
29	1460.5	80.2	2500	6	AR562011
30	1460.5	80.2	2500	6	AX333242
31	1460.5	80.2	2500	6	AX333687
32	1460.5	80.2	2500	6	AX334345
33	1460.5	80.2	2500	9	HSRR2SS
34	1456.5	80.0	3039	9	AK092671
35	1445.5	79.4	3203	9	AK123010
36	1444	79.3	1538	5	CR762018
37	1444	79.3	1611	5	CR484103
38	1433	79.2	2111	10	MUSRNRM2A
39	1439	79.0	1406	5	CR353642
40	1439	79.0	1647	5	BC080161
41	1438	79.0	21774	2	AC120603
42	1431	78.6	1605	5	BC047975
43	1429	78.5	1170	6	E14825
44	1428.5	78.4	1600	5	BC041209
45	1425.5	78.3	1328	5	DRUS7965

ALIGNMENTS

RESULT 1	AR454868	Sequence 2 from patent US 6682917.	1053 bp	DNA	linear	PAT 20-FEB-2004
LOCUS	AR454868	Sequence 2 from patent US 6682917.	1053 bp	DNA	linear	PAT 20-FEB-2004
DEFINITION	AR454868	Sequence 2 from patent US 6682917.	1053 bp	DNA	linear	PAT 20-FEB-2004
ACCESSION	AR454868	Sequence 2 from patent US 6682917.	1053 bp	DNA	linear	PAT 20-FEB-2004
VERSION	AR454868.1	GI:42688823	1053 bp	DNA	linear	PAT 20-FEB-2004
KEYWORDS	AR454868.1	GI:42688823	1053 bp	DNA	linear	PAT 20-FEB-2004
SOURCE	Unknown.	Unknown.	1053 bp	DNA	linear	PAT 20-FEB-2004
ORGANISM	Unknown.	Unknown.	1053 bp	DNA	linear	PAT 20-FEB-2004
REFERENCE	1 (bases 1 to 1053)	Unclassified.	1053 bp	DNA	linear	PAT 20-FEB-2004
AUTHORS	Nakamura, Y., Arakawa, H. and Tanaka, H.	Unclassified.	1053 bp	DNA	linear	PAT 20-FEB-2004
TITLE	Protein having a ribonucleotide Reductase activity and a DNA thereof	Unclassified.	1053 bp	DNA	linear	PAT 20-FEB-2004
JOURNAL	Patent: US 6682917-A 2 27-JAN-2004;	Unclassified.	1053 bp	DNA	linear	PAT 20-FEB-2004
FEATURES	source	Location/Qualifiers	1053 bp	DNA	linear	PAT 20-FEB-2004
ORIGIN	1..1053	/organism="unknown"	1053 bp	DNA	linear	PAT 20-FEB-2004
	/mol_type="genomic DNA"		1053 bp	DNA	linear	PAT 20-FEB-2004

Alignment Scores:	1.01e-170	Length:	1053
Pred. No.:	1821.00	Matches:	351
Score:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-698-228-1 (1-351) x AR454868 (1-1053)		Db	1021	ACAGATAACGTCTTCCACCTTGGATGCAGATTTT	1053
RESULT 2					
AR454877					
LOCUS					
DEFINITION				1053 bp	DNA
ACCESSION				AR454877	linear
VERSION				AR454877.1	GI:42688832
KEYWORDS				Unknown.	
SOURCE				Unknown.	
ORGANISM				Unclassified.	
REFERENCE				1 (bases 1 to 1053)	
AUTHORS				Nakamura, Y., Arakawa, H. and Tanaka, H.	
TITLE				Protein having a ribonucleotide Reductase activity and a DNA	
JOURNAL				Patent: US 6682917-A 12 27-JAN-2004;	
FEATURES				Location/Qualifiers	
source				1..1053	
ORIGIN				/organism="unknown"	
				/mol_type="genomic DNA"	
Alignment Scores:					
Pred. No.:				1,01e-170	Length: 1053
Score:				1821.00	Matches: 351
Percent Similarity:				100.00%	Conservative: 0
Best Local Similarity:				100.00%	Mismatches: 0
Query Match:				100.00%	Indels: 0
DB:				6	Gaps: 0
US-10-698-228-1 (1-351) x AR454877 (1-1053)					
Qy	1	MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer	20		
Db	1	ATGGGCGACCCCGAAAGCGGAGCGCGGCTGGATCAGATGAGATCATCTTCA	60		
Qy	21	AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg	40		
Db	61	GACACCAAGAAAGTGAATTAAGTCAATGAAGCCACTCTTAAGAAAGATTCTCGC	120		
Qy	41	ArgPheValIlePheProLeuArgPheSerHisLeuLeuAlaPheAsnAlaGln	60		
Db	121	CGGTTTGTCTATCTTCCATCCAGTACCTGATATTGGAAATGTATAACAGGCACAG	180		
Qy	61	AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys	80		
Db	181	GCTTCTCTGGACAGAGAGGTTGACTTATCAAGGATCTCCCTCACTGGAACAAG	240		
Qy	81	LeuLysAlaAspGluLysTyrPheIleSerHisLeuAlaPheAsnAlaSerAsp	100		
Db	241	CTTAAAGCAGATGAGAAGTACTTCTCATCTCACATCTTAGCCTTTTTCAGCCAGTGAT	300		
Qy	101	GlyIleValAsnGluAsnLeuValArgPheSerGlnGluValGlnValProGluAla	120		
Db	301	GGAAATTGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCCAGAGGCT	360		
Qy	121	ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu	140		
Db	361	CGCTGTTTCTATGGCTTTCAAAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG	420		
Qy	141	LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu	160		
Db	421	CTGATAGACATTTACATCAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA	480		
Qy	161	ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys	180		
Db	481	ACCATGCCCTATGTTAAGAAAAAGCAGATTGGCGCTTCGATGGATAGCAGATAGAAA	540		
Qy	181	SerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePheSerGly	200		
Db	541	TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGGAGTCTTCTTCACGGA	600		
Qy	201	SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer	220		
Db	601	TCTTTTGTGCTATATCTGGCTTAAAGAAAGAGGCTTATGCGAGGACTCACTTTTTC	660		
Qy	221	AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln	240		
Db	661	ANTGAATCTCATCAGCAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA	720		
Qy	241	TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys	260		
Db	721	TACTTAGTAAATAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGTGATGCTGCAAA	780		
Qy	261	IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle	280		
Db	781	ATTGACGAGAGTATTAACAGAGCCTTGCCAGTTGGCTCATTTGGAATGCAATTCAT	840		
Qy	281	LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer	300		
Db	841	TTGATGAACAGTACATGAGTTTGTAGTGCAGAGATTACTTGTGAACCTTGGATCTCA	900		
Qy	301	LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnLysSerLeuGluGlyLys	320		
Db	901	AAGGTTTTTTCAGCAGAAAATCCTTTGATTTTATGGAATAACATTCTTTTAGAAGGAAA	960		
Qy	321	ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr	340		
Db	961	ACAAATTTCTTTTGAAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAC	1020		
Qy	341	ThrAspAsnValPheThrLeuAspAlaAspPhe	351		

```

QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 601 TCTTTTGTCTATATCTGGCTAAAGAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
QY 221 AnGluLeuLysSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGTGCTTGGCTGTATGCCAA 720

QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluLileValAspAlaValLys 260
Db 721 TACTTAGTAATTAAGCTTCAGAGAAAGGGTCAGGAGATCATTTGATGCTGTCAA 780
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 781 ATTGAGCAGGAGTCTTTTAAACAGAGCTTGGCCATTGGCTCATTTGGAATGATTCATT 840
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuGlyPheSer 300
Db 841 TTGATGAACACGTACATAGTGTGTAGCTGACAGATTACTTGTGAACTTGGATTCTCA 900
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 901 AAGGTTTTTCAGCGAGAAATCTTTTGAATTTTATGGAATAACATTCTTTAGAGGAAAA 960
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 961 ACAATTTCTTTGAGNAACAGTTTCAGAGTATCAGCGTTTGCAGTTATGGCAGNAACC 1020
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1021 ACAGATAACGTCTTCACTTGGATGCAGATTTT 1053

RESULT 3
BD064764
LOCUS
DEFINITION
ACCESSION BD064764
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
OS Homo sapiens (human)
PN JP 2001269184-A/1
PD 02-OCT-2001
PF 27-JUN-2000 JP 2000192401
PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
C12N15/09, A61K31/711, A61K38/00, A61K38/43, A61K45/00, A61K48/00, PC
A61P35/00,
PC A61P43/00, C07K16/32, C07K16/40, C12N1/15, C12N1/19, C12N1/21 PC
, C12N5/10, C12N9/02,
PC C12P21/02, C12P21/08, G01N33/15, G01N33/50, G01N33/53// (C12N15/09,
PC C12R1:91),
PC (C12N15/09, C12R1:91), (C12N9/02, C12R1:91), (C12P21/02, C12R1:91),
PC (C12P21/08, C12R1:91), (C12N15/00, A61K37/02, A61K37/48, C12N5/00,
PC (C12N15/00, C12R1:91), (C12N15/00, C12R1:19)
CC Novel protein and DNA thereof
FH Key Location/Qualifiers
FT source 1..1053
/organism='Homo sapiens (human)'
/organism='Homo sapiens'
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'

FEATURES
source
ORIGIN

```

```

Alignment Scores:
Pred. No.: 1,01e-170 Length: 1053
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-698-228-1 (1-351) x BD064764 (1-1053)
QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20
Db 1 ATGGCGACCCGCGAAGAGCGCGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
QY 21 AspThrAsnGluSerGluLileLysSerAsnGluLileProLeuLeuArgLysSerSerArg 40
Db 61 GACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
Db 121 CGTTTGTCTATCTTTCCATCCAGTACCTGATATTTGGAAAATGATATAACAGGCACAG 180
QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
Db 181 GCTTCCTTCTGGACAGCAGAGAGGTGCGATTATCAAGAGATCTCCCTCACTGGAACAAG 240
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAspSerAsp 100
Db 241 CTTAAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTAGCCCTTTTTCAGCCAGTGTAT 300
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 301 GGAATTTGTAATGAAATTTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGTTCCAGAGGCT 360
QY 121 ArgCysPheTrpGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAAGTGTTCACCTCAGAGATGTACAGTTTG 420
QY 141 LeuIleAspThrTyrIleArgAspProLysGlyArgGluPheLeuPheAsnAlaIleGlu 160
Db 421 CTGATAGACACTTACATCAGATATCCAGAAAGGGAATTTTATTTAATGCAATTGAA 480
QY 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Db 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGGCATGGATAGAGATAGAAAA 540
QY 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
Db 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTAGAGAGGAGTTTCTTCTCAGGA 600
QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 601 TCTTTTGTCTATATCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
QY 221 AnGluLeuLysSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGTGCTTGGCTGTATGCCAA 720
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluLileValAspAlaValLys 260
Db 721 TACTTAGTAATTAAGCTTCAGAGAAAGGGTCAGGAGATCATTTGATGCTGTCAA 780
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 781 ATTGAGCAGGAGTCTTTTAAACAGAGCTTGGCCACTTGGCTCATTTGGAATGATTCATT 840
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuGlyPheSer 300
Db 841 TTGATGAACACGTACATAGTGTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 901 AAGGTTTTTCAGCGAGAAATCTTTTGAATTTTATGGAATAACATTCTTTAGAGGAAAA 960

```

```

QY 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 961 ACAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTGCGATTTATGGCAGAAACC 1020
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1021 ACAGATAACGCTTTCACCTTGGATGCGAGATTTT 1053

RESULT 4
BD064773
LOCUS BD064773 1053 bp DNA linear PAT 27-AUG-2002
DEFINITION Novel protein and DNA thereof.
ACCESSION BD064773
VERSION BD064773.1 GI:22610376
KEYWORDS JP 2001269184-A/10.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Nakamura,Y., Arakawa,H. and Tanaka,H.
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL 1 (bases 1 to 1053)
COMMENT YUSUKE NAKAMURA,HIROFUMI ARAKAWA,HIROSHI TANAKA PC
OS Homo sapiens (human)
PN JP 2001269184-A/10
PD 02-OCT-2001
PI 27-JUN-2000 JP 2000192401
PF YUSUKE NAKAMURA,HIROFUMI ARAKAWA,HIROSHI TANAKA PC
PC C12N15/09,A61K31/711,A61K38/00,A61K38/43,A61K45/00,A61K48/00,PC
PC A61P43/00,C07K16/32,C07K16/40,C12N1/15,C12N1/19,C12N1/21 PC
PC C12N5/10,C12N9/02,
PC C12P21/02,C12P21/08,G01N33/15,G01N33/50,G01N33/53//(C12N15/09,
PC C12R1:91),
PC (C12N15/09,C12R1:19),(C12N9/02,C12R1:91),(C12P21/02,C12R1:91),
PC (C12P21/08,C12R1:91),C12N15/00,A61K37/02,A61K37/48,C12N5/00,
PC (C12N15/00,C12R1:91),(C12N15/00,C12R1:19)
CC Novel protein and DNA thereof
FH Key Location/Qualifiers
FT source 1..1053
FT /organism='Homo sapiens (human)'.

FEATURES
source
1..1053
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 1,01e-170 Length: 1053
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-698-228-1 (1-351) x BD064773 (1-1053)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
Db 1 ATGGGGGACCCGGAAGGCGGAGCGCGGCTGGATCAGATCAGAGATCATCTTCA 60
QY 21 AspThrAsnGluSerGluLeuLysSerAsnGluGluProLeuLeuArgLysSerArg 40
Db 61 GACACCAACGAAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 120
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrPheMetTyrLysGlnAlaGln 60
Db 121 CGGTTTGTTCATCTTCCATCCAGTACCTGATATTTGGAAATGTATAAAGCAGCAG 180
QY 61 AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80

```

```

181 GCCTTCCTTCGGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAAACAAG 240
81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
241 CTTAAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTAGCCCTTTTTCGACCCAGTAT 300
101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
301 GGAATTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGAGGTCGAGGTTCCAGAGGCT 360
121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCAGAAATGTTCTACTCAGAGATGTACAGTTTG 420
141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
421 CTGATAGACATTTACATCAGATGCCAAGAAAAGGGAATTTTATTATTAAATGCAATTGAA 480
161 ThrMetProTyrValLysLysAlaAspTyrAlaLeuArgTyrPheAlaAspArgLys 180
481 ACCATGCCCTATGTTAAGAAAAGAGCATTTGGCCCTTGGATGGATAGCAGATAGAAA 540
181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGluValPhePheSerGly 200
541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGAGAGAGATTTTCTTCTCAGGA 600
201 SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
601 TCTTTTGGCTGCTATATTTCTGGCTAAAGAGAGAGGCTTATGCCAGGAGCTCCTTTTTC 660
221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
661 AATGAATCTCATCAGCAGATGAAGGACTTCACTGTGACTTTGCTGCTGATGTTTCCAA 720
241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
721 TACTTAGTAATAAGCCCTTCAGAAAGAGGCTCAGGAGAGATCATTTGTTGATGCTGTCAA 780
261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
781 ATTGAGCAGGAGTATTTTAAACAGAGCCCTTGCCAGTTCGCTCATTTGGAATGAATTCAT 840
281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
841 TTGATGAAACAGTACATTTAGTGTGTAGCTGACAGATTTACTTTGGAACTTTGGATTTCA 900
301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
901 AAGGTTTTTCAGGCAGAAATCCTTTGATTTTATGGAAAACATTTCTTTAGAGGAAAA 960
321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
961 ACAAAATTTCTTTGAGAAACAGATTTTACAGATATCAGCGTTTTCAGGTTTATGGCAGAAACC 1020
341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
1021 ACAGATAACGCTTTCACCTTGGATGCGAGATTTT 1053

RESULT 5
BD093077
LOCUS BD093077 1053 bp DNA linear PAT 27-AUG-2002
DEFINITION Novel protein and its DNA.
ACCESSION BD093077
VERSION BD093077.1 GI:22638665
KEYWORDS WO 0100799-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Nakamura,Y., Arakawa,H. and Tanaka,H.
TITLE Novel protein and its DNA

```

JOURNAL		Patent: WO 0100799-A 1 04-JAN-2001; TAKEDA CHEMICAL INDUSTRIES LTD, YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA	
COMMENT		OS Homo sapiens (human) PN WO 0100799-A/1 PD 04-JAN-2001 PF 27-JUN-2000 WO 2000JP004189 PR 28-JUN-1999 JP 99P 181131, 06-JUL-1999 JP 99P 192391 PR PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC C12N9/04, C12N15/53, C12N1/15, C12N1/19, C12N1/21, C12N5/10 PC , C12P21/02, A61K38/44, PC A61K48/00, C07K16/40, G01N33/53, C12Q1/26 CC	
FEATURES	source	Location/Qualifiers	
		1..1053 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"	
ORIGIN			
Alignment Scores:			
Pred. No.:		1,01e-170	Length:
Score:		1821.00	Matches:
Percent Similarity:		100.00%	Conservative:
Best Local Similarity:		100.00%	Mismatches:
Query Match:		100.00%	Indels:
DB:		6	Gaps:
US-10-698-228-1 (1-351) x BD093077 (1-1053)			
QY	1	MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer	20
DB	1	ATGGGCGACCCCGAAGCGCGAAGCGCGCGCTGGATCAGATGAGATCATCTTCA	60
QY	21	AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg	40
DB	61	GACACCAACGAAGTGAATAAGTCAATGAAGGCCACTCTCTAAGAAGAGTTCTCGC	120
QY	41	ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln	60
DB	121	CGTTTGTTCATCTTCCATCCAGTACCTGATATTGGAAATGATATAACAGGCACAG	180
QY	61	AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys	80
DB	181	GCTTCCTCTGACAGCAGAGAGTTCGACTTATCAAGGATCTCCCTCACTGGAACAAG	240
QY	81	LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp	100
DB	241	CTTAAAGCAGATGAGAAGTACTCTCTCATCTTTCACATCTTTCGACCCAGTGTAT	300
QY	101	GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla	120
DB	301	CGAATTTGTAATGAAATTTTGTGGAGCGCTTGTAGTCAGAGGTGCGAGGTTCAGAG	360
QY	121	ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu	140
DB	361	CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGATGTTTCACTCAGAGATGTACAGTTG	420
QY	141	LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu	160
DB	421	CTGATAGACACTTATCATGAGATCCCAAGAAAGGGAATTTTATTTTATTCATTTGAA	480
QY	161	ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys	180
DB	481	ACCATGCCCTATGTTAAGAAAAGCAGATTGGGCTTCGGATGGATGATAGAGAAA	540
QY	181	SerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePheSerGly	200
DB	541	TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAAGGAGTTTCTTCTCAGCA	600
QY	201	SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer	220

Db	601	TCITTTGCTCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACITTTTCC	660
QY	221	AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln	240
DB	661	AATGAACATCATCAGCAGAGATGAGGACTTTCACCTGTGACTTTGCTGCTGATGTCCAA	720
QY	241	TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys	260
DB	721	TACTTAGTAATAAGCCTTCAGAAAGAGGTCAAGGAGATCATTTGTTGATGCTGTCAA	780
QY	261	IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle	280
DB	781	ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTTGSCCTCATTTGAATGAATTCATT	840
QY	281	LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer	300
DB	841	TTGATGAAACAGTACATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	900
QY	301	LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys	320
DB	901	AAGTTTTCAGGCAGAAAATCCTTTTGAATTTTATGAAAACATTTCTTTTGAAGGAAA	960
QY	321	ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr	340
DB	961	ACAAATTTCTTTCAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGTCAGAAACC	1020
QY	341	ThrAspAsnValPheThrLeuAspAlaAspPhe	351
DB	1021	ACAGATAACGCTTTCACCTTGGATGCAGATTTT	1053

BD093086	Novel protein and its DNA.	1053 bp	DNA	linear	PAT 27-AUG-2002
BD093086	BD093086_1 GI-22638674				
KEYWORDS	WO 0100799-A/10.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 1053)				
AUTHORS	Nakamura, Y., Arakawa, H. and Tanaka, H.				
TITLE	Novel protein and its DNA				
JOURNAL	Patent: WO 0100799-A 10 04-JAN-2001; TAKEDA CHEMICAL INDUSTRIES LTD, YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA				
COMMENT	OS Homo sapiens (human) PN WO 0100799-A/10 PD 04-JAN-2001 PF 27-JUN-2000 WO 2000JP004189 PR 28-JUN-1999 JP 99P 181131, 06-JUL-1999 JP 99P 192391 PR PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC C12N9/04, C12N15/53, C12N1/15, C12N1/19, C12N1/21, C12N5/10 PC , C12P21/02, A61K38/44, PC A61K48/00, C07K16/40, G01N33/53, C12Q1/26 CC				
FEATURES	source	Location/Qualifiers			
		1..1053 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"			
ORIGIN					
Alignment Scores:					
Pred. No.:		1,01e-170	Length:	1053	
Score:		1821.00	Matches:	351	
Percent Similarity:		100.00%	Conservative:	0	
Best Local Similarity:		100.00%	Mismatches:	0	
Query Match:		100.00%	Indels:	0	
DB:		6	Gaps:	0	

DB:	6	Gaps:	0
US-10-698-228-1 (1-351) x BD093086 (1-1053)			
QY	1	MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer	20
DB	1	ATGGGCGACCCCGGAAGCGCGGCGCTGGATCAGATGAGAGATCATCTTCA	60
QY	21	AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg	40
DB	61	GACACCAACCAAGTGAATAAGTCAATGAAGACCACTCTTAAGAAGAGTTCTCGC	120
QY	41	ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln	60
DB	121	CGGTTTGTTCATCTTCCATCCAGTACCTGATATTTGGAAATGTATAAACAGGCACAG	180
QY	61	AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys	80
DB	181	GCTTCTCTTCTGGACAGCAGAGAGGTGACTTATCAAGGATCTCCCTCACTGGAACAAG	240
QY	81	LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaIleSerAsp	100
DB	241	CTTAAGCAGATGAGAAGTACTTCATCTCTACATCTTAGCCCTTTTTCAGCCAGTGAT	300
QY	101	GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnProGluAla	120
DB	301	CGAATTGTAATGAATAATTTGGTGGAGCGCTTAGTCAGAGGTGCAGGTTCCAGAGGCT	360
QY	121	ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu	140
DB	361	CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAAGTGTTCACCTCAGAGATGTACAGTTTG	420
QY	141	LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu	160
DB	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAAATGCAATTGAA	480
QY	161	ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys	180
DB	481	ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTGGCATGATAGCAGATAGAAA	540
QY	181	SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly	200
DB	541	TCTACTTTTGGGGAAGAGTGGTGGCTTTCGTCTGTAGAAGAGTTCCTTCTCAGGA	600
QY	201	SerPheAlaAlaIlePheThrProLysLysArgGlyLeuMetProGlyLeuThrPheSer	220
DB	601	TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC	660
QY	221	AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln	240
DB	661	AATGAACCTCATCAGCAGATGAGGACTTCATGTGACTTTTGTGCTGATGTTCCAA	720
QY	241	TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys	260
DB	721	TACTTAGTAATAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGTCATGCTGTCAAA	780
QY	261	IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle	280
DB	781	ATTGAGCAGGAGTGTATTAACAGAAAGCTTGGCCAGTGGCTCATTTGGAAATGAATTCAT	840
QY	281	LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer	300
DB	841	TTGATGAACAGTACATTGAGTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTTCTCA	900
QY	301	LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys	320
DB	901	AAGGTTTTTCAGGCAGAAAATCCTTTGATTTATGGAATAACATTTCTTTAGAAGGAAA	960
QY	321	ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr	340
DB	961	ACAAATTTCTTTGAGAAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC	1020
QY	341	ThrAspAsnValPheThrLeuAspAlaAspPhe	351

DB	1021	ACAGATAACGCTCTTCCCTTGGATGACAGATTTT	1053
RESULT 7			
LOCUS	CQ714252	1056 bp	DNA
DEFINITION	Sequence 186 from Patent WO02068579.		linear
ACCESSION	CQ714252		
VERSION	CQ714252.1	GI:42275109	
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens		
REFERENCE	1		
AUTHORS	Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.		
TITLE	Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof		
JOURNAL	Patent: WO 02068579-A 186 06-SEP-2002;		
FEATURES	PE Corporation (NY) (US)		
source	Location/Qualifiers		
	1..1056		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Alignment Scores:			
Pred. No.:	1,01e-170	Length:	1056
Score:	1821.00	Matches:	351
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-10-698-228-1 (1-351) x CQ714252 (1-1056)			
QY	1	MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer	20
DB	1	ATGGGCGACCCCGGAAGCGCGGCTGGATCAGATGAGAGATCATCTTCA	60
QY	21	AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg	40
DB	61	GACACCAACCAAGTGAATAAGTCAATGAAGACCACTCTTAAGAAGAGTTCCTCGC	120
QY	41	ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln	60
DB	121	CGGTTTGTTCATCTTCCATCCAGTACCTGATATTTGGAAATGTATAAACAGGCACAG	180
QY	61	AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys	80
DB	181	GCTTCTCTTGGACAGCAGAGAGGTGACTTATCAAGGATCTCCCTCACTGGAACAAG	240
QY	81	LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaIleSerAsp	100
DB	241	CTTAAGCAGATGAGAAGTACTTCATCTCTACATCTTAGCCCTTTTTCAGCCAGTGAT	300
QY	101	GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnProGluAla	120
DB	301	CGAATTGTAATGAATAATTTGGTGGAGCGCTTAGTCAGAGGTGCAGGTTCCAGAGGCT	360
QY	121	ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu	140
DB	361	CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAAGTGTTCACCTCAGAGATGTACAGTTTG	420
QY	141	LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu	160
DB	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAAATGCAATTGAA	480
QY	161	ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys	180
DB	481	ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTGGCATGATAGCAGATAGAAA	540

```

Qy 181 SerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePheSerGly 200
Db 541 TCTACTTTTGGGAAGAAGTGTGGCTTGTCTGTAGAGAGATTTCTTCTCAGGA 600

Qy 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 601 TCTTTTGTCTGTATATTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTCC 660

Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 661 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720

Qy 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleValAspAlaValLys 260
Db 721 TACTTAGTAATAAGCTTCAGAGAAAGGGTCAGGAGATCATTTGATGCTGCAAA 780

Qy 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 781 ATTGACGAGGAGTCTTTTAAACAGAGCTTGCCAGTTGGCTCATTTGAATGATTCATT 840

Qy 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 841 TTGATGAACAGTACATTCAGTTGTAGCTGACAGATTAATTTGTGAACTTGGATTCTCA 900

Qy 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 901 AAGGTTTTTCAGGCGAGAAATCCCTTTGATTTATGGAAGAAACATTTCTTTAGAAGAAA 960

Qy 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 961 ACAATTTCTTTGAGAACGAGTTTCAGAGTATCAGCGTTTGCAGTTATGCGAAGAAC 1020

Qy 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1021 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1053

RESULT 8
AB166671
LOCUS
DEFINITION
Homo sapiens p53R2 mRNA for p53-inducible ribonucleotide reductase
small subunit 2, complete cds.
ACCESSION
AB166671
VERSION
AB166671.1 GI:45259568
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
Ugai,H. and Yokoyama,K.K.
Homo sapiens p53R2 mRNA for p53-inducible ribonucleotide reductase
small subunit 2, complete cds
Published Only in Database (2004)
REFERENCE
2
(bases 1 to 1056)
Ugai,H. and Yokoyama,K.K.
Direct Submission
Submitted (05-MAR-2004) Hideyo Ugai, RIKEN BioResource Center, Gene
Engineering Division; 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074,
Japan (E-mail:ugai@rtc.riken.go.jp, Tel:81-29-836-3612,
Fax:81-29-836-9120)
FEATURES
Location/Qualifiers
1..1056
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="p53R2.C19"
/cell_lines="Saos-2"
/cell_type="Osteogenic sarcoma"
1..1056
/gene="p53R2"
1..1056
/gene="p53R2"

```

```

/codon_start=1
/product="p53-inducible ribonucleotide reductase small
subunit 2"
/protein_id="BAD12267.1"
/db_xref="GI:45259568"
/translation="MGDPPEPAAGLQDSSDTNESEIKSNEEP LLKSSRRFVI
FPIQYDPIWMYKQAQASFTABEVDLSKDLPHWNKADKDYFISHLAFPAASDGI
VNEVLVERFQVQVPEARCFYGOIILINVHSEMYSLIDTYRDPKRFLEFNAIE
TWPYKKADMALRWIADRSKTRGVVFAAAGVEGFPFSGSFAALFWLKKGLMPLGT
FNEHLSRDEGLHCDFAFLMFQILVNKPSERVREIIVDAKIQEFTEALPVGLIG
MNCILMKQYIEFVADRLVLELFGFSKVFQAFNPFDFMENISLEGKTNFFEKRVSEYORF
AVMAETTDNVFTLDAF"
ORIGIN
Alignment Scores: 1.01e-170 Length: 1056
Pred. No.: 1821.00 Matches: 351
Score: 1821.00
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
DB: 9
US-10-698-228-1 (1-351) x AB166671 (1-1056)
Qy 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
Db 1 ATGGCGACCCCGGAAGCCGCGAGCGCGGCTGGATCAGATGAGATCATCTTCA 60
Qy 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
Db 61 GACACCAACGAAAGTGAATAAAGTCAATCAAGACGACCTCTTAAGAAAGAGTTCTCGC 120
Qy 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
Db 121 CGGTTTGTCTATCTTCCATCCAGTACCTCATATTGGAAAAATGTATAAACAGGCACAG 180
Qy 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
Db 181 GCTTCCTCTTGGACAGCAGAAAGAGTGCAGTTATCAAGAGGATCTCCCTCCTCGGAACAG 240
Qy 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Db 241 CTTAAGCAGATGAGAGTACTTCTCATCTCTACATCTTAGCTTTTTCGAGCCAGTAT 300
Qy 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 301 GGAATTTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGT 360
Qy 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 361 CGCTGTTTCTATGGCTTCAAAATTTCTCATCAGAAATGTTCTCATCAGATGTACAGTTTG 420
Qy 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 421 CTGATAGACACTTACATCAGATCCCAAGAAAGGAAATTTTATTATTATGCAATTGAA 480
Qy 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Db 481 ACCATGCCCTATCTTAAGAAAAAAGCAGATTGGGCCCTTGCATGATGATGAGATAGAAAA 540
Qy 181 SerThrPheGlyGluArgValAlaPheAlaValGluGlyValPhePheSerGly 200
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCTTGTCTGTAGAGAGATTTCTTCTCAGGA 600
Qy 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 601 TCTTTTGTCTGTATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTCC 660
Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 661 AATGAACATCATCAGCAGAGATGAAGACTTTCACGTGACTTTTGTGCTGATGTTCCAA 720
Qy 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleValAspAlaValLys 260

```

Db 721 TACTTAGTAAATAAGCTTCAGAAAGAGGGTCAGGAGATCATTTGATGCTGTCAAA 780
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 Db 781 ATGTAGCAGAGAGTTTAAACAGAGCTTGGCAGTTGGCTCATTTGGAATGAATTGCATT 840
 QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 Db 841 TTGATGAAACAGTACATTGAGTTGTAGCTGACAGATTACTTGTGAACTTGGATTCTCA 900
 QY 301 LysValPheGluAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 Db 901 AAGGTTTTTCAGCGCAAAAATCCTTTGATTTTATGCAAAACATTTCTTTAGAGGAAAA 960
 QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 Db 961 ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGATTATGCGAAGAAC 1020
 QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 Db 1021 ACAGATAAGCTCTTCACCTTGGATGAGATTTT 1053

RESULT 9

AR454870

LOCUS AR454870 1081 bp DNA linear PAT 20-FEB-2004

DEFINITION Sequence 4 from patent US 6682917.

ACCESSION AR454870

VERSION AR454870.1 GI:42688825

KEYWORDS

SOURCE Unknown.

ORGANISM

Unclassified.

REFERENCE 1 (bases 1 to 1081)

AUTHORS Nakamura,Y., Arakawa,H. and Tanaka,H.

TITLE Protein having a ribonucleotide Reductase activity and a DNA

JOURNAL Patent: US 6682917-A 4 27-JAN-2004;

FEATURES

source

Location/Qualifiers

1..1081

/organism="unknown"

/mol_type="genomic DNA"

ORIGIN

Alignment Scores:
 Pred. No.: 1,04e-170 Length: 1081
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-698-228-1 (1-351) x AR454870 (1-1081)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
 Db 20 ATGGCGCACCCGGAAGCGGAGCGCGGGCTGGATCAGATGAGATCATCTTCA 79
 QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
 Db 80 GACACCAACGAAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 139
 QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrPheMetTyrLysGlnAlaGln 60
 Db 140 CGGTTTGTGTCATCTTCCATCCAGTACCCTGATATTTGGAAAAATGATATAACAGGCACAG 199
 QY 61 AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisIlePheLys 80
 Db 200 GCTTCTCTCGACAGCAGAGAGGTGACTTATCAAGAGATCTCCCTCAGTGGAAACAG 259
 QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp 100
 Db 260 CTTAAGCAGATGAGAAGTACTTCTCATCTCTTACATCTTAGCCTTTTTCAGCCAGTGAT 319

QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 Db 320 GGAATTTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAGAGGTCGAGGTCAGAGGCT 379
 QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 Db 380 CGCTGTTTCTATGGCTTTCAAATTTCTCATCAGAAATGTTCTCACTCAGAGATGACAGTTTG 439
 QY 141 LeuIleAspThrTyrIleArgAspProLysValArgGluPheLeuPheAsnAlaIleGlu 160
 Db 440 CTGATAGACACTTACATCAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA 499
 QY 161 ThrMetProTyrValLysLysAlaAspTrioAlaLeuArgTrioIleAlaAspArgLys 180
 Db 500 ACCATGCCCTATGTTTAAAGAAAAAGCAGATTGGCCCTTCGATGATAGCAGATAGAAA 559
 QY 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
 Db 560 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGTAGAAAGGAGTTTCTTCTCAGGA 619
 QY 201 SerPheAlaAlaIlePheThrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 Db 620 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGCTTATGCGAGACTCACATTTTTC 679
 QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 Db 680 ATGAACTCATCAGCAGAGATGAGAGCTTCACGTGACTTTGCTTGCCTGATGTTCCA 739
 QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
 Db 740 TACTTAGTAAATAAGCCTTCAGAAAGAAAGGCTCAGGAGATCATTTGTTGCTGTCAA 799
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 Db 800 ATTGACGAGAGTTTTTAACAGAAAGCCTTGCAGATTGGCCTCATTTGGAATGAATGCA 859
 QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 Db 860 TTGATGAAACAGTACATTTGAGTTTGTAGCTGACAGATTACTTTGGAACTTGGATTCTCA 919
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 Db 920 AAGTTTTTTCAGGCAGAAAATCCTTTTGAATTTATGAAAACATTTCTTTAGAAAGGAAA 979
 QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 Db 980 ACAAATTTCTTTCAGAAACAGATTTTCAGAGTATCAGCGTTTTTGCAGTTATGCGAAGAAC 1039
 QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 Db 1040 ACAGATAAGCTCTTCACCTTGGATGCAATTTT 1072

RESULT 10

BD064766

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 1081)

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

Nakamura, Y., Arakawa, H. and Tanaka, H.

Novel protein and DNA thereof

Patent: JP 2001269184-A 3 02-OCT-2001;

YUSUKE NAKAMURA, TAKEDA CHEMICAL INDUSTRIES LTD

OS Homo sapiens (human)

PN JP 2001269184-A/3

PD 02-OCT-2001

PF 27-JUN-2000 JP 2000192401

BD064766 1081 bp DNA linear PAT 27-AUG-2002
 Novel protein and DNA thereof.

BD064766

BD064766.1 GI:22610369

JP 2001269184-A/3.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

1 (bases 1 to 1081)

Nakamura, Y., Arakawa, H. and Tanaka, H.

Novel protein and DNA thereof

Patent: JP 2001269184-A 3 02-OCT-2001;

YUSUKE NAKAMURA, TAKEDA CHEMICAL INDUSTRIES LTD

OS Homo sapiens (human)

PN JP 2001269184-A/3

PD 02-OCT-2001

PF 27-JUN-2000 JP 2000192401


```
US-10-698-228-1 (1-351) x BD093079 (1-1081)
QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20
Db 20 ATGGCGCAGCCCGAAGCGCGGAGCGCGGCTGGATCAGAGATGATCATCTTCA 79
QY 21 AspThrAsnGluSerGlnLeuLeuSerAsnGluGluProLeuLeuArgLysSerArg 40
Db 80 GACACCAACGAAAGTAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 139
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrPheMetTyrLysGlnAlaGln 60
Db 140 CGGTTTGTATCTTCCCAATCCAGTACCTGATATTTGGAAATGTATATAACAGGCACAG 199
QY 61 AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
Db 200 GCTTCTCTTCGACAGCAGAGAGGTGACTTATCAAGAGATCTCCCTCAGTGGAAACAAG 259
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Db 260 CTTAAGCAGATGAGAGTACTTCTCATCTCAGCTTACGCTTTTTCGACCCAGTGAT 319
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 320 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 379
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 380 CGCTGTTTCTATGGCTTTCAAATTCATCGAGAATGTTCTCAGAGATGTACAGTTTG 439
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 440 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 499
QY 161 ThrMetProTyrValLysLysLysAlaAspTyrPheAlaLeuArgTyrPheAlaAspArgLys 180
Db 500 ACCATGCCCTATGTTAAGAAAGACAGATTGGGCTTCGATGATAGCAGATAGAAAA 559
QY 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
Db 560 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGTAGAGAGGATTTCTTCTCAGGA 619
QY 201 SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 620 TCTTTTGTCTATATTTCTGGCTAAGAAAGAGAGGTCTTATGCCAGGACTCACTTTTCC 679
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 680 AATGAACCTCATCAGCAGAGATGAAGSACTTCTACTGTGACTTTGCTTGCCTGATGTTCAA 739
QY 241 TyrLeuValAsnLysProSerGluArgValArgGluIleIleValAspAlaValLys 260
Db 740 TACTTTAGTAAATAACCTTCAGAAGAAAGGGTCAGGAGATCATTTGTTGATGCTGTCAA 799
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 800 ATTGAGCAGGAGTTTTTAACAGAGCCTTGCAGTTGGCTCATTTGGGAATGAAATGCAAT 859
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 860 TTGATGAAACAGTACATTTGAGTTGTAGCTGACAGATTTACTTGTGGAACCTTGATTTCA 919
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 920 AAGGTTTTTTCAGGCAGAAAAATCCTTTGATTTATTTAGAAAAACATTTCTTTAGAGGAAA 979
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 980 ACAAAATTTCTTTGAAACAGAGTTTCAGAGATATCAGCGTTTGGCAGTTATGGCAGAAACC 1039
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
|||||
```

```
Db 1040 ACAGATAACGCTCTTTCACCTTGGATGCAGATTTT 1072
RESULT 12
AR454869
LOCUS DEFINITION
AR454869 Sequence 3 from patent US 6682917. DNA linear PAT 20-FEB-2004
ACCESSION AR454869
VERSION AR454869.1 GI:42688824
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4955)
AUTHORS Nakamura,Y., Arakawa,H. and Tanaka,H.
TITLE Protein having a ribonucleotide Reductase activity and a DNA
JOURNAL thereof
FEATURES Patent: US 6682917-A 3 27-JAN-2004;
Location/Qualifiers
source 1..4955
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores: 7.39e-170 Length: 4955
Pred. No.: 1821.00 Matches: 351
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 6
US-10-698-228-1 (1-351) x AR454869 (1-4955)
QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20
Db 245 ATGGCGCAGCCCGAAGCGCGGAGCGCGGCTGGATCAGAGATGATCATCTTCA 304
QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
Db 305 GACACCAACGAAAGTAAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 364
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysGlnAlaGln 60
Db 365 CGGTTTGTCTATCTTTCCAAATCCAGTACCTTGATATTTGGAAATGTATAACAGGCACAG 424
QY 61 AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
Db 425 GCTTCTCTTCTGGACAGAGAGGTGCGACTTATCAAGAGATCTCCCTCAGTGGAAACAAG 484
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Db 485 CTTAAGCAGATGAGAGTACTTCACTCTCAGATCTTACCTTTTTCGACCCAGTGAT 544
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 545 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 604
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 605 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAATGTTCTCACTCAGAGATGTACAGTTTG 664
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 724
QY 161 ThrMetProTyrValLysLysLysAlaAspTyrPheAlaLeuArgTyrPheAlaAspArgLys 180
Db 725 ACCATGCCCTATGTTAAGAAAGACAGATTTGGCCTTGGATGATGATGATGATGATA 784
QY 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
Db 785 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAAAGAGATTTTCTTCTCAGGA 844
```

QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 Db 845 TCTTTTGTCTATATCTGGCTAAAGAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 904
 QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 Db 905 AATGAACATCATCAGCAGAGATGAAGACTTCACCTGTGACTTTGCTTGCCTGATGTCCAA 964
 QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
 Db 965 TACTTAGTAATAAGCTTCAGNAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAAA 1024
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 Db 1025 ATTGACGAGGAGTTTTTAACAGAGGCTTGGCAGTTGGCTCATTTGGAATGCAATT 1084
 QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 Db 1085 TTGATGAACACAGTACATTCAGTTTGTAGCTGACAGATTAATCTGTGAACTTGGATTCTCA 1144
 QY 301 LysValPheGluAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 Db 1145 AAGTTTTCAGCAGAGAAATCCTTTGATTTTATGGAAGAAATTTCTTTAGAGGAGAAA 1204
 QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 Db 1205 ACAAAATTCCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTGCAGTTATGGCAGAAACC 1264
 QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 Db 1265 ACAGATAACGCTCTCACCTTGGATGCAGATTTT 1297

RESULT 13
 BD064765
 LOCUS
 DEFINITION Novel protein and DNA thereof.
 ACCESSION BD064765
 VERSION BD064765.1 GI:22610368
 KEYWORDS JP 2001269184-A/2.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Nakamura, Y., Arakawa, H. and Tanaka, H.
 TITLE Novel protein and DNA thereof
 JOURNAL Patent: JP 2001269184-A 2 02-OCT-2001;
 YUSUKE NAKAMURA, TAKEDA CHEMICAL INDUSTRIES LTD
 COMMENT OS Homo sapiens (human)
 PN JP 2001269184-A/2
 PD 02-OCT-2001
 PF 27-JUN-2000 JP 2000192401
 PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
 C12N15/09, A61K31/711, A61K38/00, A61K38/43, A61K45/00, A61K48/00, PC
 A61P35/00,
 PC A61P43/00, C07K16/32, C07K16/40, C12N1/15, C12N1/19, C12N1/21 PC
 C12N5/10, C12N9/02,
 PC C12P21/02, C12P21/08, G01N33/15, G01N33/50, G01N33/53/// (C12N15/09,
 PC C12R1:91),
 PC (C12N15/09, C12R1:91), (C12N9/02, C12R1:91), (C12P21/02, C12R1:91),
 PC (C12P21/08, C12R1:91), C12N15/00, A61K37/02, A61K37/48, C12N5/00,
 PC (C12N15/00, C12R1:91), (C12N15/00, C12R1:19)
 CC Novel protein and DNA thereof
 FH Key Location/Qualifiers
 FT source 1. .4955
 /organism='Homo sapiens (human)'.
 Location/Qualifiers
 1. .4955
 /organism='Homo sapiens'
 /mol_type='genomic DNA'
 /db_xref='taxon:9606'

FEATURES
 source
 ORIGIN

Alignment Scores:
 Pred. No.: 7,39e-170 Length: 4955
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-698-228-1 (1-351) x BD064765 (1-4955)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20
 Db 245 ATGGCGCACCCGGAAGCGCGGCGGTGGATCAGGATCAGAGATCATCTTCA 304
 QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
 Db 305 GACACCAACGAAAGTGAATAAGTCAATGAAGAGCCACTCCTTAAGAAAGAGTCTCGC 364
 QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
 Db 365 CGGTTTGTCTATCTTTCCATCCAGTACCTGATATTTGGAANAATGTATAACAGGCACAG 424
 QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
 Db 425 GCTTCCCTTCTGGACAGCAGAGAGGTGCACCTTATCAAAGGATCTCCCTCACTGGAACAAG 484
 QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
 Db 485 CTTAAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTTAGCCTTTTTCAGCCAGTGTAT 544
 QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 Db 545 GGAATTTGTAATGAAATTTGGTGAGCGCTTTAGTCAGAGAGTGCAGGTTCCAGAGGCT 604
 QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 Db 605 CGCTGTTTCTATGCTTTCAAATTTCTCATCAGAAATGTTCACTCAGAGATGTACAGTTTG 664
 QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAAATCAATTTGAA 724
 QY 161 ThrMetProTyrValLysLysLeuAlaAspTrpAlaLeuArgTyrIleAlaAspArgLys 180
 Db 725 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGGCTTTGCCGATGATAGCAGATAGAAA 784
 QY 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
 Db 785 TCTACTTTTGGGAAAGAGTGGTGGCCTTTGCTGCTGTAGAGAGGATTTTCTTCTCAGGA 844
 QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 Db 845 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 904
 QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 Db 905 AATGAACATCATCAGCAGAGATGAAGACTTCACCTGTGACTTTGCTTGTGCTGATGTCCAA 964
 QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
 Db 965 TACTTAGTAATAAGCTTCAGNAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAAA 1024
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 Db 1025 ATTGACGAGGAGTTTTTAACAGAGGCTTGGCAGTTGGCCTCATTTGGAATGAATTCATT 1084
 QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 Db 1085 TTGATGAACACAGTACATTCAGTTTGTAGCTGACAGATTAATCTGTGAACTTGGATTCTCA 1144
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 Db 1145 AAGTTTTCAGCAGAGAAATCCTTTGATTTTATGGAAGAAATTTCTTTAGAGGAGAAA 1204

Qy	321	ThrAsnPhenheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr	340
Db	1205	ACAAATTCCTTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTGTATGGCAGAAACC	1264
Qy	341	ThrAsnValPheThrLeuAspAlaAspPhe	351
Db	1265	ACAGATAACGTCCTTCACCTTGGATGCAGATTTT	1297
RESULT 14			
BD093078			
LOCUS	BD093078	4955 bp	DNA linear PAT 27-AUG-2002
DEFINITION	Novel protein and its DNA.		
ACCESSION	BD093078		
VERSION	BD093078.1	GI:22638666	
KEYWORDS	WO 0100799-A/2.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Nakamura, Y., Arakawa, H. and Tanaka, H.		
TITLE	Novel protein and its DNA		
JOURNAL	Patent: WO 0100799-A 2 04-JAN-2001;		
	TAKEDA CHEMICAL INDUSTRIES LTD, YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA		
COMMENT	OS Homo sapiens (human) PN WO 0100799-A/2 PD 04-JAN-2001 PF 27-JUN-2000 WO 2000JP004189 PR 28-JUN-1999 JP 99P 181131, 06-JUL-1999 JP 99P 192391 PR 21-JAN-2000 JP 00P 017770 PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC C12N9/04, C12N15/53, C12N1/15, C12N1/19, C12N1/21, C12N5/10 PC C12P21/02, A61K38/44, PC A61K48/00, C07K16/40, G01N33/53, C12Q1/26 CC		
FEATURES	Key	Location/Qualifiers	
source	1..4955		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Alignment Scores:			
Pred. No.:	7,39e-170	Length:	4955
Score:	1821.00	Matches:	351
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-10-698-228-1 (1-351) x BD093078 (1-4955)			
Qy	1	MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer	20
Db	245	ATGGCGACCCGGAAGCCGGAAGCCGGCGCTGGATCAGGATGAGATCATCTTCA	304
Qy	21	AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg	40
Db	305	GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC	364
Qy	41	AtcPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln	60
Db	365	CGGTTTGTCACTTTTCAATCCAGTACCCCTGAATTTGGAAAAATGTATAACAGGCACAG	424
Qy	61	AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys	80
Db	425	GCTTCTCTTGGACAGCAGAGAGGTCGACTTATCAAGAGGATCTCCCTCTACTGGACAG	484
Qy	81	LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp	100

REFERENCE 2 (bases 1 to 4955)
 AUTHORS Tanaka.H., Arakawa.H. and Nakamura.Y.
 TITLE Direct Submission
 JOURNAL Submitted (18-DEC-1999) Yusuke Nakamura, University of Tokyo,
 Institute of Medical Science, Human Genome Center, Laboratory of
 Molecular Medicine; 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639,
 Japan [E-mail:yusuke@ims.u-tokyo.ac.jp. Tel:+81-3-5449-5372,
 Fax:+81-3-5449-5433]

FEATURES

Source
 1..4955
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="8"
 /map="8q23.1"
 1..4955
 /gene="p53R2"
 245..1300
 /gene="p53R2"
 /map="8q23.1"
 /codon_start=1
 /product="ribonucleotide reductase"
 /protein_id="BA92434.1"
 /db_xref="GI:7229086"
 /translation="MGDPERPEAAGLDQDERSSDTNESEIKSNBEPLLRKSSRFVI
 FPQYDWMKYQQAQSWTAEVDLSKOLPHWNKLKADKRYFISHILAFPAASDGI
 VNENIYERFSQSVQPEARCFYGFQILLENVHSEMYSLIDITYIRDPKRELFNAIE
 TNPYKKRDAWLRWIDRKSTFGERVAFPAAVEGFFSGSFAAIFWLKRGIMPGILT
 FSNELISDEGHDCDFACILFYQILVKNKPSERREIIVDAVKIEQFLEALPVGILIG
 NCILMKQYIEVFADRLIVELGFSKVQFQENPDPFMENTISLEGKTNFFPKRVSQRF
 AVMAETTONVFLDADF"
 4933
 /gene="p53R2"
 /note="22 a nucleotides"

polya_site

ORIGIN

Alignment Scores:
 Pred. No.: 7,39e-170 Length: 4955
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0
 US-10-698-228-1 (1-351) x AB036063 (1-4955)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20
 Db 245 ATGGCGACCCGGAAAGCCGGAAGCGCGCGGCTGGATCAGAGATCATCTTCA 304
 QY 21 AspThrAsnGluSerGluLeuIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
 Db 305 GACACCAACGAAAGTGAATAAAGTCAATGAGAGCCACTCTTAAGAAAGAGTTCTCGC 364
 QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
 Db 365 CGGTTTGTCATCTTCCATCCAGTACCTCTGATATTGGAAATGATTAACAGGCACAG 424
 QY 61 AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
 Db 425 GCTTCTCTCTGGACAGCAGAGAGTGCAGTATCAAGGATCTCCCTCACTGGAAACAAG 484
 QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
 Db 485 CTTAAAGCAGATGAGAAGTACTTCTATCTCATCTTAGCCCTTTTTCAGCCAGTGAT 544
 QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 Db 545 GGAATTTGAAATGAAATTTTGGTGAGCGCTTTAGTCAGAGAGTGCGAGGCTTCCAGAGGCT 604
 QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 Db 605 CGCTGTTCTATGGCTTTCANATTCTCATCGAGAATGTTCTCATCAGAGATGTACAGTTTG 664

QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTATTATGCAATTCAA 724
 QY 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
 Db 725 ACCATGCCCTATGTTTAAAGAAAAAAGCAGATTTGGGCCCTTGGCATGCATAGCAGATAGAAA 784
 QY 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
 Db 785 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAAGGAGTTTTCTTCTCAGGA 844
 QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 Db 845 TCTTTTGGCTGCTATATTTCTGGCTAAAGAAGAGAGGTCTTATGCCAGGACTCCTTTTTC 904
 QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 Db 905 AATGAATCTCATCAGCAGATGAAGGACTTCACTGTGCTTGTGCTTGTGCTGCTGCTGCTCAA 964
 QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
 Db 965 TACTTAGTAAATAAGCCTTCAGAAAGAAAGGTCAGGGAGATCATTTGTGATGCTGTCAA 1024
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 Db 1025 ATTGACGAGAGTTTTTTTAAACAGAACCTTCCAGATTGGCTTCATTTGGAATGAATGCATT 1084
 QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuGlyPheSer 300
 Db 1085 TTGATGAACAGTACATTTAGTTTGTAGCTGACAGATTACTTTGTGGAACCTTGGATTCTCA 1144
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 Db 1145 AAGGTTTTTTCAGGCAGAAAAATCCTTTTGTATTTATGAAAAACATTTCTTTAGAAGGAAAA 1204
 QY 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 Db 1205 ACAAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTTGCAGTTATGCGAGAAACC 1264
 QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 Db 1265 ACAGATAACGCTCTTCACCTTGGATGCAGATTTT 1297

Search completed: October 30, 2005, 08:09:50
 Job time : 3484 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 30, 2005, 06:58:12 ; Search time 443 Seconds
(without alignments)
4690.359 Million cell updates/sec

Title: US-10-698-228-1

Perfect score: 1821

Sequence: 1 MGDPERPBAAGLDQDERSSS.....QRFVMAETTDNVFTLDADF 351

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q/cgn2.1/USPTO.spool/US10698228/runat 26102005.100714.5097/app query.fasta 1.519
-DB=N Geneseq 16Dec04 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pcio -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10698228 @CGN 1.1.352 @runat 26102005.100714.5097 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 16Dec04.*

1: Geneseqn1980s.*

2: Geneseqn1990s.*

3: Geneseqn2000s.*

4: Geneseqn2001as.*

5: Geneseqn2001bs.*

6: Geneseqn2002as.*

7: Geneseqn2002bs.*

8: Geneseqn2003as.*

9: Geneseqn2003bs.*

10: Geneseqn2003cs.*

11: Geneseqn2003ds.*

12: Geneseqn2004as.*

13: Geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1821	100.0	1053	4	AAF32438	Aaf32438 Human rib
2	1821	100.0	1053	4	AAF32447	Aaf32447 Human rib
3	1821	100.0	1081	4	AAF32440	Aaf32440 Human rib
4	1821	100.0	2596	3	AAAI2411	Aaai2411 cDNA enco
5	1821	100.0	4955	4	AAF32439	Aaf32439 Human rib

6	1821	100.0	4955	13	ADR24210	Adr24210 Breast ca
7	1817	99.8	1601	4	AAH14924	Aah14924 Human cdn
8	1460.5	80.2	1989	3	AAC78111	Aac78111 Human can
9	1460.5	80.2	2215	12	ADK70302	Adk70302 Respiraco
10	1460.5	80.2	2216	10	ADJ56536	Adj56536 Murine cd
11	1460.5	80.2	2482	4	AAS44917	Aas44917 Human con
12	1460.5	80.2	2500	6	ABL65414	Abi65414 Lung can
13	1460.5	80.2	2500	6	ABL65517	Abi65517 Lung can
14	1460.5	80.2	2500	6	ABL65859	Abi65859 Lung can
15	1460.5	80.2	2500	8	ABX10335	Abx10335 DNA encod
16	1460.5	80.2	2500	11	ADI32044	Adi32044 Human cdn
17	1460.5	80.2	2500	12	ADNO4443	Adno4443 Antipsori
18	1460.5	80.2	2500	12	ADQ09273	Adq09273 Human RRM
19	1460.5	80.2	2500	13	ACN37637	Acn37637 Tumour-as
20	1460.5	80.2	3294	4	AAH73225	Aah73225 Human cer
21	1460.5	80.2	3393	12	ADN03788	Adn03788 Antipsori
22	1460.5	80.2	3393	12	ADO19225	Ado19225 Human PRO
23	1447.5	79.5	2641	4	AAS44745	Aas44745 Human ful
24	1429	78.5	1170	2	AAV05641	Aav05641 Human rib
25	1425.5	78.3	1328	12	ADO57308	Ado57308 DNA encod
26	1376.5	75.6	1371	5	AAS79474	Aas79474 DNA encod
27	1348	74.0	1289	4	ABL14627	Abi14627 Drosophil
28	1241.5	68.2	3945	4	ABL14626	Abi14626 Drosophil
29	1230	67.5	1218	8	ABT20895	Abt20895 Aspergill
30	1220.5	67.0	1292	13	ADS49379	Ads49379 Bacterial
31	1210	66.4	14176	2	AAT84564	Aat84564 Swinepox
32	1210	66.4	14176	4	AAF84949	Aaf84949 Reverse c
33	1210	66.4	14176	4	AAF84948	Aaf84948 Nucleotid
34	1192.5	65.5	1146	8	ABT19075	Abt19075 Aspergill
35	1192	65.5	1242	6	ABZ32250	Abz32250 Candida a
36	1188	65.2	706	13	ADQ57092	Adq57092 Novel can
37	1176.5	64.6	963	13	ADS58378	Ads58378 Bacterial
38	1168.5	64.2	1200	10	ACC61238	Acc61238 Gene sequ
39	1168.5	64.2	1200	10	ADK63241	Adk63241 Disease t
40	1168.5	64.2	1200	13	ADT47419	Adt47419 Bacterial
41	1161.5	63.8	1173	13	ADT47814	Adt47814 Bacterial
42	1161.5	63.8	1206	10	ADB69740	Adb69740 C. neofo
43	1151	63.2	1314	8	ABT18481	Abt18481 Aspergill
44	1151	63.2	1450	8	ABT20297	Abt20297 Aspergill
45	1151	63.2	3314	8	ABT17887	Abt17887 Aspergill

ALIGNMENTS

RESULT 1
AAF32438
ID AAF32438 standard; cdna; 1053 BP.
XX
XX
AC AAF32438;
XX
DT 18-APR-2001 (first entry)
XX
DE Human ribonucleotide reductase TP53R2H nucleotide sequence SEQ ID NO:2.
XX
KW Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.
XX
OS Homo sapiens.
XX
PN WO200100799-A1.
XX
PD 04-JAN-2001.
XX
PF 27-JUN-2000; 2000WO-JP004189.
XX
PR 28-JUN-1999; 99JP-00181131.
PR 06-JUL-1999; 99JP-00192391.
PR 21-JAN-2000; 2000JP-00017770.
XX
(TAKE) TAKEDA CHEM IND LTD.
(NAKA/) NAKAMURA Y.
XX
PI Nakamura Y, Arakawa H, Tanaka H;
XX

DR WPI; 2001-112446/12.
 XX P-PSDB; AAB69050.
 PT Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
 XX diagnosis, treatment and prevention of cancer.
 XX Claim 5; Fig 1-3; 102pp; Japanese.
 PS
 CC The present sequence encodes a human ribonucleotide reductase designated
 CC TP53R2H. TP53R2H has cytostatic activity. TP53R2H is part of the DNA
 CC repair mechanism and its activity is induced by p53. It can be used for
 CC the treatment, prevention and diagnosis of a wide range of cancers
 XX
 SQ Sequence 1053 BP; 314 A; 192 C; 246 G; 301 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 4,17e-212 Length: 1053
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-10-698-228-1 (1-351) x AAF32438 (1-1053)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
 DB 1 ATGGGGCAGCCCGAAGGCGGAGCGGCGGCTGGATCAGGATGAGATCATCTTCA 60
 QY 21 AspThrAsnGluSerGluLeuLysSerAsnGluProLeuLeuArgLysSerSerArg 40
 DB 61 GACACCAACGAAAGTCAATTAAGTCAATGAAGAGCCATCTCTAAGAAAGAGTCTCGC 120
 QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
 DB 121 CGGTTTGTCTATCTTCCATCCAGTACCTGCTGATATTGGAAAATGATATAACAGGCACAG 180
 QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
 DB 181 GCTTCCTCTGGACACAGAGAGGTCGATCAATCAAGGATCTCCCTACTCGAACAAG 240
 QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAlaSerAsp 100
 DB 241 CTTAAGCAGATGAGAAGTACTTTCATCTCTCATCTTAGCCTTTTTCGAGCCAGTGAT 300
 QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 DB 301 CGAATTGTAAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGAGGCT 360
 QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 DB 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTG 420
 QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 DB 421 CTGATAGACATACATCAGATCCAGAAAGGGAATTTTATTAATGCAATTGAA 480
 QY 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
 DB 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGCCCTTCGATGGATAGCAGATAGAAA 540
 QY 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
 DB 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAGAGAGTCTTCTCTCAGGA 600
 QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 DB 601 TCTTTTGTCTATATCTGGCTTAAGAAAGAGAGGTCCTTATGCCAGAGCTCACTTTTTC 660
 QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 DB 661 AATGAACATCATCAGCAGATGAAGACCTTCACTGTGACTTTGCTTGGCTGATGTTCCAA 720

QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
 DB 721 TACTTAGTAATAAGCCTTCAGAGAAGAGGTCAAGGAGATCATTTGATGCTGTCAA 780
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 DB 781 ATTGAGCAGGAGTATTTAAACAGAACGCTTGCCAGTTGGCCTCATTTGGAATGAATGCATT 840
 QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 DB 841 TTGATGAAAACAGTACATTTGATGTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTTCA 900
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 DB 901 AAGGTTTTTCAGGCAGAAATCTTTGATTTTATGGAAACATTTCTTTAGAGGAAAA 960
 QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 DB 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGATTATGCGAGAAACC 1020
 QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 DB 1021 ACAGATAACGTCCTTCACCTTGGATGCAGATTTT 1053
 RESULT 2
 AAF32447
 ID AAF32447 standard; cDNA; 1053 BP.
 XX
 AC AAF32447;
 XX
 DT 18-APR-2001 (first entry)
 XX
 DE Human ribonucleotide reductase related nucleotide sequence SEQ ID NO:12.
 XX
 KW Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200100799-A1.
 XX
 PD 04-JAN-2001.
 XX
 PF 27-JUN-2000; 2000WO-JP004189.
 XX
 PR 28-JUN-1999; 99JP-00181131.
 PR 06-JUL-1999; 99JP-00192391.
 PR 21-JAN-2000; 2000JP-00017770.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 PA (NAKA/) NAKAMURA Y.
 XX
 PI Nakamura Y, Arakawa H, Tanaka H;
 XX
 DR WPI; 2001-112446/12.
 XX
 PT Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
 PT diagnosis, treatment and prevention of cancer.
 XX
 PS Claim 6; Page 96-97; 102pp; Japanese.
 XX
 CC The present invention describes a human ribonucleotide reductase
 CC designated TP53R2H. TP53R2H has cytostatic activity. TP53R2H is part of
 CC the DNA repair mechanism and its activity is induced by p53. It can be
 CC used for the treatment, prevention and diagnosis of a wide range of
 CC cancers. The present sequence represents a human ribonucleotide reductase
 CC related sequence which is given in the exemplification of the present
 CC invention
 XX
 SQ Sequence 1053 BP; 314 A; 191 C; 246 G; 302 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 4,17e-212 Length: 1053
 Score: 1821.00 Matches: 351

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-698-228-1 (1-351) x AAF32447 (1-1053)

```
QY 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSer 20
DB 1 ATGGGCGACCGGAAAGCGCGAGCGCGCGGCTGGATCAGATCAGATCATCTTCA 60
QY 21 AspThrAsnGluSerGluLeuIleYssAsnGluGluProLeuLeuArgLysSerArg 40
DB 61 GACACCAACGAAAGTGAATAAGTCAATGAAGAGCCACTCTAAGAAAGAGTTCTCGC 120
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrPheMetTyrLysGlnAlaGln 60
DB 121 CGGTTTGTTCATCTTCCATCCAGTACCTGATATTTGGAAATGTATTAACAGGCACAG 180
QY 61 AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
DB 181 GTTCTCTCTGACAGCAGAGAGGTGCTTATCAAGGATCTCCCTCCTGGAACAAG 240
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaIleSerAsp 100
DB 241 CTTAAAGCAGATGAGAAGTACTTCTCTCATCTTACATCTTGGCTTTTTCGAGCCAGTAT 300
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 301 GGAATTTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 361 CGCTGTTTCTATGGCTTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGACAGTTTG 420
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 421 CTGATAGACACTTATCATCAGATCCCAAGAAAGGAAATTTTATTAATGCATTTGAA 480
QY 161 ThrMetProTyrValLysLysLysAlaAspTyrPheAlaLeuArgTyrPheAlaAspArgLys 180
DB 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGCGCTTCGATGGATAGCAGATAGAAA 540
QY 181 SerThrPheGlyGluArgValAlaPheAlaValGluGlyValPhePheSerGly 200
DB 541 TCTACTTTTGGGGAAGAGTGTGGCTTGTCTGTAGAGGAGTTTCTTCTCAGA 600
QY 201 SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 601 TCTTTTGTCTATATTCTGGCTAAGAAAGAGAGGTTCTATGCCAGACTCATTTTTC 660
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
DB 661 AATGAACCTCATCAGCAGACATCAGAGCTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
DB 721 TACTTAGTAAATTAAGCTTTCAGAAAGAAAGGGTCAGGGAGATCATTTGTTGATGCTGCAAA 780
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
DB 781 ATTGACGAGGAGTTTAAACAGAGCCTTGCAGTTGGCTCATTTGGAATGAATTCATT 840
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
DB 841 TTGATGAAACACTACATTCAGTTTGTAGCTGACAGATTAATCTTGTGAACTTGGATTTCTCA 900
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
DB 901 AAGGTTTTTTCAGCGCAGAAATCTTTTGAATTTATGTAAGAAACATTTCTTTTAGAAGGAAA 960
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
```

```
DB 961 ACAAATTTCTTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTATGCAGAAACC 1020
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB 1021 ACAGATAACGTCCTTCACCTTGGATGCAGATTTT 1053
```

RESULT 3

```
AAF32440
ID AAF32440 standard; cDNA; 1081 BP.
XX AAF32440;
DT 18-APR-2001 (first entry)
XX Human ribonucleotide reductase related nucleotide sequence SEQ ID NO:4.
XX Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.
XX Homo sapiens.
XX WO200100799-A1.
XX 04-JAN-2001.
XX 27-JUN-2000; 2000WO-JP004189.
XX 28-JUN-1999; 99JP-00181131.
PR 06-JUL-1999; 99JP-00192391.
PR 21-JAN-2000; 2000JP-00017770.
XX (TAKE ) TAKEDA CHEM IND LTD.
PA (NAKA/) NAKAMURA Y.
XX Nakamura Y, Arakawa H, Tanaka H;
PI WPI; 2001-112446/12.
XX Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
PT diagnosis, treatment and prevention of cancer.
PS Example 2; Page 91; 102pp; Japanese.
XX The present invention describes a human ribonucleotide reductase
CC designated TP53R2H. TP53R2H has cytostatic activity. TP53R2H is part of
CC the DNA repair mechanism and its activity is induced by p53. It can be
CC used for the treatment, prevention and diagnosis of a wide range of
CC cancers. The present sequence represents a human ribonucleotide reductase
CC related sequence which is used in an example from the present invention
XX
SQ Sequence 1081 BP; 322 A; 199 C; 252 G; 308 T; 0 U; 0 Other;
```

Alignment Scores:

Pred. No.:	4,34e-212	Length:	1081
Score:	1821.00	Matches:	351
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-10-698-228-1 (1-351) x AAF32440 (1-1081)

```
QY 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSer 20
DB 20 ATGGGCGACCGGAAAGCGCGAGCGCGGCTGGATCAGATGAGATCATCTTCA 79
QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
DB 80 GACACCAACGAAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 139
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrPheMetTyrLysGlnAlaGln 60
DB 140 CGGTTTGTCTATCTTCCATCCAGTACCTTGATTTGGAAATGTATTAACAGGCACAG 199
```

QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
 DB 200 GCTTCCTCTGACAGCAGAGAGGTTGACTTATCAAGGATCTCCTCTCACTGGACAAG 259
 QY 81 LeuLysAlaAspGluLysTyrrPheIleSerHisIleLeuAlaPhePheAlaIleAsp 100
 DB 260 CTTAAAGCAGATGAGAGTACTCTCTCATCTCATCTTTCAGCCAGTGTAT 319
 QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 DB 320 GGAATTTGAATGAATAATTTGGTGAGCGCTTTAGTCAGAGGTGCAGGTTTCAGAGGCT 379
 QY 121 ArgCysPheTyrrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrrSerLeu 140
 DB 380 CGCTGTTTCTATGGCTTTCAAAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 439
 QY 141 LeuIleAspThrTyrrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 DB 440 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 499
 QY 161 ThrMetProTyrrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
 DB 500 ACCATGCCCTATGTTAAGAAAAGCAGATTCGGCTTCGATGGATGATCAGATAGAAA 559
 QY 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
 DB 560 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTGTAGAGGAGTTTCTTCACGA 619
 QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 DB 620 TCTTTTGTCTGTATATCTGGCTAAAGAGAGAGGTTCTTATGCCAGACTCACTTTTTC 679
 QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 DB 680 AATGAACCTATCAGCAGAGATCAGAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 739
 QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
 DB 740 TACTTAGTAAATAAGCTTCAGAGNAAGGTCAGGAGATCATTTGTGATGCTGCMAA 799
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 DB 800 ATTGAGCAGGAGTTTAAACAGAAGCCTTGCCAGTGTGCTCATTGGAATGAATTCATT 859
 QY 281 LeuMetLysGlnTyrrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 DB 860 TTGATGAACACAGTACATTGAGTTGTAGTCAGAGATTACTTGTGGAATTTGGATTCTCA 919
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 DB 920 AAGGTTTTTCAGCGAGAAATCCTTTTGATTTTATGGAACAATTTCTTTAGAGGAAA 979
 QY 321 ThrAsnPheGluLysArgValSerGluTyrrGlnArgPheAlaValMetAlaGluThr 340
 DB 980 ACAAAATTTCTTTGAAACACAGTTTCAGAGTATCAGCGTTTTCAGCTTATGGCAGAAACC 1039
 QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 DB 1040 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1072

RESULT 4
 AAA12411
 ID AAA12411 standard; cDNA; 2596 BP.

XX AAA12411;

DT 25-JUL-2000 (first entry)

XX cDNA encoding a human RNA-associated protein.

XX Human; RNA-associated protein; cell proliferation; cancer; inflammation;
 KW immune response; reproductive disorder; actinic keratosis;
 KW atherosclerosis; arteriosclerosis; bursitis; cirrhosis; hepatitis;

KW mixed connective tissue disease; myelofibrosis; primary thrombocythemia;
 KW paroxysmal nocturnal hemoglobinuria; polycythemia vera; psoriasis;
 KW trauma; ss.

XX Homo sapiens.

OS Location/Qualifiers

XX 56..1111

FT /*tag= a

FT /product= "RNA-associated protein"

XX WO200015799-A2.

PN 23-MAR-2000.

XX 17-SEP-1999; 99WO-US021688.

XX 17-SEP-1998; 98US-00156039.

PR 22-SEP-1998; 98US-00158720.

PR 04-NOV-1998; 98US-00186815.

PR 08-APR-1999; 99US-0128660P.

XX (INCY-) INCYTE PHARM INC.

XX Tang YT, Corley NC, Guegler KJ, Gorgone GA, Patterson C;

PI Hillman JL, Baughn MR, Lal P, Azimzai Y, Yue H, Yang J;

XX WPI; 2000-271437/23.

DR P-PSDB; AAY84439.

XX New polypeptides and polynucleotides, useful for preventing and treating

PT a disorder associated with increased or decreased expression of RNA

PT associated proteins.

PS Claim 9; Page 120-121; 131pp; English.

XX The present sequence encodes a human RNA-associated protein. The

CC expression of RNA-associated proteins is closely associated with

CC reproductive tissues, nervous tissues, cell proliferation including

CC cancer, inflammation and immune responses, and so they may be used for

CC diagnosis, treatment or prevention of cell proliferative,

CC immune/inflammatory disorders, and reproductive disorders. Diseases and

CC disorders which may be treated include actinic keratosis,

CC atherosclerosis, arteriosclerosis, bursitis, cirrhosis, hepatitis, mixed

CC connective tissue disease, myelofibrosis, paroxysmal nocturnal

CC hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia

CC and cancers, and trauma

XX SQ Sequence 2596 BP; 783 A; 418 C; 510 G; 885 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.59e-211 Length: 2596

Score: 1821.00 Matches: 351

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 3 Gaps: 0

US-10-698-228-1 (1-351) x AAA12411 (1-2596)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20

DB 56 ATGGGCGACCCGGAAGAGCGCGCGGCTGGATCAGATCAGAGATCATCTTCA 115

QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40

DB 116 GACACCAACGAAAGTGAATAAAGTCAATGAAGACCACTCTTAAGAAAGAGTTCTCGC 175

QY 41 ArgPheValIlePheProIleGlnTyrrProAspIleTrpLysMetTyrrLysGlnAlaGln 60

DB 176 CGGTTTGTCTATCTTTCATCCAGTACCTCATATTTGGAAATGATAAAGCAGCAG 235

QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80

```
Db 236 GCTTCCTCTGACAGCAGAGAGGTGCGACTTATCAAGGATCTCCTCACTGGAACAAG 295
Qy 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaIleSerAsp 100
Db 296 CTTAAAGCAGATGAGAAGTACTCTCATCTCATCTTAGCCCTTTTTCAGCCAGTGAT 355
Qy 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 356 GGAATGTAAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 415
Qy 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 416 CGCTGTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTG 475
Qy 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 476 CTGATAGACACTTACATCAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA 535
Qy 161 ThrMetProTyrValLysLysValAlaAspTTPAlaLeuArgTyrPheAlaAspArgLys 180
Db 536 ACATGCCCTATGTTAAGAAAAGCAGATTGGGCTTCGGATGGATGAGATAGAGAAA 595
Qy 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
Db 596 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGA 655
Qy 201 SerPheAlaAlaIlePheTyrPheLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 656 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGCTTTATGCCAGGACTCACTTTTTC 715
Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 716 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTGTGCTGATGTTCCAA 775
Qy 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
Db 776 TACTTAGTAAATAGCCTTCAGAGNAGGGTTCAGGAGATCATTTGTTGATGCTGTCAAA 835
Qy 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 836 ATTGAGCAGGAGTCTTTTAAACAGAGCCTTGCAGTTGGCTCATTTGGAATGAATTCATT 895
Qy 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 896 TTGATGAAACACATACATTGAGTTTGTAGCTGACAGATTACTTTGTGGAACCTTGATTC 955
Qy 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 956 AAGGTTTTTTCAGCGAGAAAATCCTTTTGATTTTATGGAACAATTTCTTTAGAAGAAA 1015
Qy 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 1016 ACAAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGATTATGCAGAAACC 1075
Qy 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1076 ACAGATAAGCTCTTCACCTTGGATGAGATTTT 1108
```

RESULT 5

AAF32439 standard; cDNA; 4955 BP.

XX AC AAF32439;

XX DT 18-APR-2001 (first entry)

XX DE Human ribonucleotide reductase related nucleotide sequence SEQ ID NO:3.

XX KW Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.

XX OS Homo sapiens.

```
PN WO200100799-A1.
XX 04-JAN-2001.
XX 27-JUN-2000; 2000WO-JP004189.
XX 28-JUN-1999; 99JP-00181131.
PR 06-JUL-1999; 99JP-00192391.
XX 21-JAN-2000; 2000JP-00017770.
XX (TAKEDA ) TAKEDA CHEM IND LTD.
PA (NAKA/) NAKAMURA Y.
XX Nakamura Y, Arakawa H, Tanaka H;
XX WPI; 2001-112446/12.
DR PT Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
PT diagnosis, treatment and prevention of cancer.
XX PS Example 2; Page 87-90; 102pp; Japanese.
XX CC The present invention describes a human ribonucleotide reductase
CC designated TP53R2H. TP53R2H has cytostatic activity. TP53R2H is part of
CC the DNA repair mechanism and its activity is induced by p53. It can be
CC used for the treatment, prevention and diagnosis of a wide range of
CC cancers. The present sequence represents a human ribonucleotide reductase
CC related sequence which is used in an example from the present invention
XX
XX Sequence 4955 BP; 1526 A; 803 C; 999 G; 1627 T; 0 U; 0 Other;
```

Alignment Scores:

Pred. No.: 4,13e-211 Length: 4955
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-698-228-1 (1-351) x AAF32439 (1-4955)

```
Qy 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20
Db 245 ATGGGCGACCCGGAAAGCCGAGCGCGCGCTGGATCAGGATGAGAGATCATCTTCA 304
Qy 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
Db 305 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 364
Qy 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
Db 365 CGGTTTGTTCATCTTTCATATCCAGTACCTGATATTTGGAAAAATGTATAACAGGCACAG 424
Qy 61 AlaSerPheTyrPheAlaGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
Db 425 GCTTCTCTTGGACAGCAGAGAGGTCGACTTATCAAGAGGATCTCCCTCACTGGAACAAG 484
Qy 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Db 485 CTTAAAGCAGATGAGAAGTACTCTCATCTCTCATCTTAGCCCTTTTTCAGCCAGTGAT 544
Qy 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 545 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 604
Qy 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 605 CGCTGTTTCTATGGCTTTCAAAATTCATCAGAAATGTTCTCAGAGATGTACAGTTG 664
Qy 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 665 CTGATAGACACTTACATCAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA 724
```

QY 161 ThrMetProTyrValLysLysLysAlaAspTrrPalaLeuArgTrpIleAlaAspArgLys 180
 Db |||||
 725 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGCCCTTCGGATGGATAGACAGATAGAAA 784
 QY 181 SerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePheSerGly 200
 Db |||||
 785 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGTGTAGAAAGAGTCTTCTTCAGGA 844
 QY 201 SerPheAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 Db |||||
 845 TCTTTTGGCTGTATTCCTGGCTTAAGAAAGAGGCTTATGCCAGAGCTCACTTTTCC 904
 QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 Db |||||
 905 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 964
 QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleValleValAspAlaValLys 260
 Db |||||
 965 TACTTAGTAATAAGCCTTCAGAAGAAAGGTCAGGGAGATCATTTGATGCTGTCAA 1024
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 Db |||||
 1025 ATTGACGAGAGTATTTTAAACAGAGCCTTGCAGTTGGCCTCATTTGGAATGAATTGCATT 1084
 QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 Db |||||
 1085 TTGATGAACAGTACATTCAGTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTTCTCA 1144
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 Db |||||
 1145 AAGGTTTTTTCAGGCAGAAAATCCTTTTGTATTTTATGGAAAAACATTTCTTTTGAAGGAAAA 1204
 QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 Db |||||
 1205 ACAAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGATTATGGCAGAAACC 1264
 QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 Db |||||
 1265 ACAGATAAGCTTTTACCTTGGATGCAGATTTT 1297

RESULT 6
 ID ADR24210 standard; DNA; 4955 BP.
 XX AC ADR24210;
 XX DT 21-OCT-2004 (first entry)
 XX DE Breast cancer prognosis marker #71.
 XX KW ds; breast cancer; prognosis; gene expression; diagnosis.
 XX OS Homo sapiens.
 XX PN WO2004065545-A2.
 XX PD 05-AUG-2004.
 XX PF 15-JAN-2004; 2004WO-US001100.
 XX PR 15-JAN-2003; 2003US-00342887.
 XX PA (ROSE-) ROSETTA INPHARMATICS LLC.
 XX PA (NECA-) NETHERLANDS CANCER INST.
 XX PI Van't Veer LJ, He Y;
 XX DR WPI; 2004-593473/57.

Classifying a breast cancer patient according to prognosis comprises
 PT determining the similarity between the level of expression of each of
 PT five genes in a cell sample taken from patient, to control levels.

PS Disclosure; SEQ ID NO 71; 226pp; English.
 XX
 CC The invention relates to a method of classifying a breast cancer patient
 CC according to prognosis by determining the similarity between the level of
 CC expression of each of five genes for which markers are listed in the
 CC specification, in a cell sample taken from the breast cancer patient, to
 CC control levels of expression for each respective five genes to obtain a
 CC patient similarity value. The methods are useful for classifying a breast
 CC cancer patient according to prognosis. Kits and computer program products
 CC are useful for data analysis using the diagnostic, prognostic and
 CC statistical methods of the invention. This sequence corresponds to a
 CC marker used in the method of the invention.
 XX

SQ Sequence 4955 BP; 1526 A; 803 C; 999 G; 1627 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4.13e-211 Length: 4955
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 13 Gaps: 0

US-10-698-228-1 (1-351) x ADR24210 (1-4955)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20
 Db |||||
 245 ATGGCGACCCCGAAGAGCGCCGGCTGGATCAGATGAGAGATCATCTCA 304
 QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
 Db |||||
 305 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 364
 QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
 Db |||||
 365 CGGTTTGTCTACTTTCATCCAGTACCTGATATTTGGAAAATGTATAAACAGGCACAG 424
 QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
 Db |||||
 425 GCTTCCTCTTGGACAGCAGAGAGGTGCTTATCAAGAGATCTCCCTCACTGGAACAG 484
 QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAspAsp 100
 Db |||||
 485 CTTAAGCAGATGAGAAGTACTTCACTCTCTCACATCTTAGCCTTTTTCAGCCAGTGAT 544
 QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 Db |||||
 545 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGGTTCCAGAGCT 604
 QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 Db |||||
 605 CGCTGTTTCTATGGCTTTCAAAATTCATCAGAAATGTTTCACTCAGAGATGATACAGTTTG 664
 QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 Db |||||
 665 CTGATAGACACTTACATCAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 724
 QY 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
 Db |||||
 725 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGCCCTTGGCATGCATGACAGATAGAAAA 784
 QY 181 SerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePheSerGly 200
 Db |||||
 785 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTAGAAAGAGTCTTCTTCTCAGA 844
 QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 Db |||||
 845 TCTTTTGTGCTATATTTCTGGCTAAGAAAGAGGCTCTTATGCCAGAGCTCACTTTTCC 904
 QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 Db |||||
 905 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 964

Qy	241	TyrIeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys	260
Db	965	TACTTAGTAATAAGCCTTCAGAGAAGGGTCAGGGAGATCATTTGTGATGCTGTCAA	1024
Qy	261	IleGluGlnGluPheIeuThrGluAlaLeuProValGlyIleuIleGlyMetAsnCysIle	280
Db	1025	ATTGAGCAGGAGTTTTAAACAGAAGCCTTGCCAGTTGGCTCTCATTTGGAATGCAATTGCATT	1084
Qy	281	LeuMetLysGlnTyrlleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer	300
Db	1085	TTGATGAAACAGATACATTGATGTTGTAGCTGACAGATTACTTGTGGAACTTGGATTCTCA	1144
Qy	301	LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys	320
Db	1145	AAGGTTTTTCAGGCAGAGAAATCCTTTTGATTTTATGGAAAACATTTCTTTAGAGGAAAA	1204
Qy	321	ThrAsnPhePheGluLysArgValSerGluTyrglnArgPheAlaValMetAlaGluThr	340
Db	1205	ACAAATTTCTTTGAGAAACAGTTTTACAGATATCAGCGTTTTCAGTTATGGCAGAAACC	1264
Qy	341	ThrAspAsnValPheThrLeuAspAlaAspPhe	351
Db	1265	ACAGATAACCTCTTCACCTTTGGATGACAGATTTT	1297

RESIST, T 7

AAH14924
ID AAH14924 standard; cDNA; 1601 BP.
XX
XX
XX
XX AC
XX AC
XX
DT 26-JUN-2001 (first entry)
XX
XX
DE Human cDNA sequence SEQ ID NO:12810.
XX
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
XX
XX 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
XX
PA (HELI-) HELIX RES INST.
XX
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.

Primer sets for synthesizing polynucleotides, particularly the 5802 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

PS Claim 8; SEQ ID NO 12810; 2537pp + Sequence Listing; English.

The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a

QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluLeuValAspAlaValLys 260
 DB 760 TACTTAGTAATAAGCTTCAGAGAAAGGTCAGGAGATCAATGTTGATGCTGTCAA 819
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuLeuGlyMetAsnCysIle 280
 DB 820 ATTGAGCAGGAGTGTATTAACAGAAAGCTTGCCAGTTGGCCCTCATTTGGAGTGAATTGCATT 879
 QY 281 LeuMetLysGluTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 DB 880 TTGATGAACAGTACATGAGTTTGAGCTGACAGATTACTTGGAACTTGGATTCTCA 939
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 DB 940 AAGGTTTTCAGGCAGAAATCTTTTGATTTATGGAAACATTTCTTTAGAGGAAA 999
 QY 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 DB 1000 ACAAAATTTCTTCAGAAACAGATTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1059
 QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 DB 1060 ACAGATAACGTTCTACCTTGGATGAGATTTT 1092

RESULT 8

AAC78111

ID AAC78111 standard; cDNA; 1989 BP.

AC AAC78111;

XX

XX

DT 08-FEB-2001 (first entry)

XX

DE Human cancer associated gene sequence SEQ ID NO:505.

XX

XX

KW Human; cancer associated gene; cancer antigen; detection; cancer;

KW diagnosis; cytostatic; proliferative; vulnerable; immunomodulator;

KW anti-diabetic; anti-asthmatic; anti-rheumatic; anti-arthritis; antiviral;

KW anti-inflammatory; antithyroid; antiallergic; antibacterial; cardiac;

KW dermatological; neuroprotective; thrombolytic; coagulant; neutropenic;

KW vasotropic; antiproliferative; antiangiogenic; gene therapy; inflammation;

KW immune disorder; haematopoietic cell disorder; autoimmune disorder;

KW allergic reaction; graft versus host disease; organ rejection;

KW haemostatic; thrombolytic; cardiovascular disorder; infection;

KW neurological disease; drug screening; ss.

XX

OS Homo sapiens.

XX

PN WO20005350-A1.

XX

XX

PD 21-SEP-2000.

XX

XX

PF 08-MAR-2000; 2000WO-US005882.

XX

PR 12-MAR-1999; 99US-0124270P.

XX

XX

PA (HUMA-) HUMAN GENOME SCI INC.

XX

XX

PI Rosen CA, Ruben SM;

XX

XX

XX

DR WPI; 2000-587533/55.

XX

DR P-PSDB; AAB43902.

XX

XX

PT Novel isolated nucleic acids comprising sequences encoding peptides

XX

PT useful for treating or diagnosing e.g. cancer.

XX

CC anti-inflammatory; antithyroid; antiallergic; antibacterial; antiviral;
 CC dermatological; neuroprotective; cardiac; thrombolytic; coagulant;
 CC neutropenic; vasotropic; antiproliferative; anti-rheumatic; anti-arthritis;
 CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of
 CC the present invention
 XX

SQ Sequence 1989 BP; 525 A; 439 C; 476 G; 544 T; 0 U; 5 Other;

Alignment Scores:

Pred. No.:	1-21e-167	Length:	1989
Score:	1460.50	Matches:	277
Percent Similarity:	88.42%	Conservative:	36
Best Local Similarity:	78.25%	Mismatches:	35
Query Match:	80.20%	Indels:	6
DB:	3	Gaps:	1

US-10-698-228-1 (1-351) x AAC78111 (1-1989)

QY 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
 DB 185 CTTGAGCGGACCGCGCTGCTGGCCAGCAAGCGGAGGAGGATCTTCCAGAGCCAC 244
 QY 23 nGluSerGluLysSer-----AsnGluGluProLeuLeuArgLys 38
 DB 245 GGAGCCGAAACTAAAGCAGCTGCCCGCGGTGGAGGATGAGCGCTGCTGAGAGAAA 304
 QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysG 58
 DB 305 CCCCCCGCGCTTTGTCATCTTCCCATCGAGTACCATGATCTGCGAGATGATTAAGA 364
 QY 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
 DB 365 GGCAGAGCTTCTTTTGGACCGCGGAGGAGGTGGACCTCTCCAAGGACATTCAGCAC 424
 QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAla 98
 DB 425 GGAATCCCTGAAACCCGAGGAGAGATATTTATATCCCATGTTCTTGGCTTTCTTGCAGC 484
 QY 98 aSerAspGlyIleValAsnGluAenLeuValGluArgPheSerGlnGluValGlnValPr 118
 DB 485 AAGCGATGGCATAGTAATGAAACTTGTGGAGCGATTTAGCCAAAGATTCAGATTAC 544
 QY 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuLeuGluAsnValHisSerGluMetTy 138
 DB 545 AGAAGCCCGCTGTTCTTCTGCTTCCAAATTCCTGATGGAACATACATCTCTGAATGA 604
 QY 138 rSerLeuLeuLeuAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
 DB 605 TAGTCTTCTTATGACACTTACATAAAAGATCCCAAGAAAGGAATTTCTTCAATGC 664
 QY 158 aIleGluThrMetProTyrValLysLysValAlaAspTrpAlaLeuArgTrpIleAlaAs 178
 DB 665 CATTGAAACCATGCTTGTGTCAAGAAAGAGCGAGACTGGGCGCTTGCCTGATGGGGA 724
 QY 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePh 198
 DB 725 CAAAGAGCTACCTATGTTGTAACGTTGTAGCTTTGCTGAGTGGAGGATTTCTTCTT 784
 QY 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
 DB 785 TTCCGGTCTTTTTCGTCGATATTCTGGCTCAAGAAACAGGAGCTGATGCTGCTGCCTCAC 844

Claim 1; Page 1035-1036; 2352pp; English.
 AAC77607 to AAC78448 encode the human cancer associated proteins given in
 AAB43398 to AAB44239. The proteins can have activities based on the
 tissues and cells the genes are expressed in. Example of activities
 include: cytostatic; proliferative; vulnerable; immunomodulator;
 anti-diabetic; anti-asthmatic; anti-rheumatic; anti-arthritis;

QY 218 rPheSerAenGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
 Db 845 ATTTTCTAATGAATTTATTAGCAGAGATGAGGGTTTACACTGTGATTTTGTCTGCTGAT 904
 QY 238 tPheGlnTyrLeuValAenLysProSerGluGluArgValArgGluIleValAspAl 258
 Db 905 GTTCAACACCTGGTACCAACCAATCGGAGGAGAGAGTAAGAGAAATAATTATCAATGC 964
 QY 258 aValIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
 Db 965 TGTTCGATAGAACAGGAGTTCTACTGAGGCTTGGCTGTGAACCTATTGGGATGA 1024
 QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuG1 298
 Db 1025 TTGCACCTAATGAAGCAATACATTGAGTTTGTGGCAGACAGACTTATCTGGAATGGG 1084
 QY 298 yPheSerLysValPheGlnAlaGluAenProPheAspPheMetGluAenIleSerLeuG1 318
 Db 1085 TTTTACCAAGGTTTTCAGAGTAGAGAACCAATTGACTTTATGGAGAAATATTCTACTGGA 1144
 QY 318 uGlyLysThrAenPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetal 338
 Db 1145 AGGAAGACTAATCTTTTGAAGAGAGTAGCGGAGTATCAGAGATCGGAGTGATGTC 1204
 QY 338 aGluThrThrAspAenValPheThrLeuAspAlaAspPhe 351
 Db 1205 AGTCCACACAGAAATCTTTTACCTTGGATGCTGACTTC 1244

RESULT 9

ADK70302
 ID ADK70302 standard; cDNA; 2215 BP.
 AC ADK70302;
 XX
 DT 06-MAY-2004 (first entry)
 DE Respiratory disease differentially expressed cDNA #38.
 XX
 KW ds; gene; cytostatic; respiratory; antiasthmatic; gene therapy;
 KW differential gene expression; respiratory disorder; lung cancer;
 KW chronic obstructive pulmonary disease; emphysema; asthma.
 XX

OS Homo sapiens.

XX WO2003101293-A2.

XX PD 11-DEC-2003.

XX PF 02-JUN-2003; 2003WO-US017409.

XX PR 04-JUN-2002; 2002US-0386003P.

XX PA (INCY-) INCYTE CORP.

XX PI Rickert PK, Krasnow R;

XX DR WPI; 2004-042945/04.

XX New combination comprising cDNAs and proteins that are differentially
 PT expressed in respiratory disorders, useful for diagnosing or treating
 PT respiratory diseases e.g. lung cancer, chronic obstructive pulmonary
 PT diseases or asthma.

XX PS Claim 1; SEQ ID NO 38; 343pp; English.

XX The invention relates to cDNA sequences that are differentially expressed
 CC in respiratory disorders or their complements or encoded proteins. The
 CC cDNAs and proteins are useful for diagnosing, treating or monitoring
 CC treatment of a subject with a respiratory disease including lung cancer,
 CC chronic obstructive pulmonary diseases, emphysema or asthma. The protein
 CC is also useful for screening molecules or compounds to identify at least
 CC one ligand which specifically binds the protein. It is also useful for
 CC preparing and purifying a polyclonal or monoclonal antibody. This

CC sequence corresponds to a cDNA of the invention.
 XX
 SQ Sequence 2215 BP; 593 A; 476 C; 523 G; 623 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1,42e-167 Length: 2215
 Score: 1460.50 Matches: 277
 Percent Similarity: 88.42% Conservative: 36
 Best Local Similarity: 78.25% Mismatches: 35
 Query Match: 80.20% Indels: 6
 DB: 12 Gaps: 1

US-10-698-228-1 (1-351) x ADK70302 (1-2215)

QY 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
 Db 188 CCTGAGCGGACCGCGCTCTGGCCAGCAAGACCGCGAGGAGGATCTTCCAGGAGCCAC 247
 QY 23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
 Db 248 GGAGCCGAAACCTAAGCAGCTGCCCGCGGTGGAGATGAGCGCTGCTGAGAGAAA 307
 QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysG1 58
 Db 308 CCCCCGCGCTTGTTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGATAAGAA 367
 QY 58 nAlaGlnAlaSerPheThrThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
 Db 368 GGCAGAGGCTTCTTTTGGACCGCGAGGAGGTGGACCTCTCCAAGGACATTCAGCACTG 427
 QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAl 98
 Db 428 GGAATCCCTGAAACCCGAGGAGAGATATTTATATCCCATGTTCTTGGCTTTCTTTCGAGC 487
 QY 98 aSerAspGlyIleValAenGluAenLeuValGluArgPheSerGlnGluValGlnValPr 118
 Db 488 AAGCGATGGCATAGTAATAAGAACTTGTGGAGCGATTAGCCAGAGATTCAGATTAC 547
 QY 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAenValHisSerGluMetTy 138
 Db 548 AGAAGCCCGCTGTTTCTATGGCTTCCAAATTTGCCATGCAAAACATACATCTCTGAAATGTA 607
 QY 138 rSerLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
 Db 608 TAGTCTTCTTATTGACACTTACATAAAAGATCCCAAGAAAGGGAATTTCTTCTCAATGC 667
 QY 158 aIleGluThrMetProTyrValLysLysAlaAspTyrAlaLeuArgTyrIleAlaAs 178
 Db 668 CATTGAACGATGCTTGTGTCAAGAGAGGCGAGACTGGGCTTGGCTGGATTGGGA 727
 QY 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePh 198
 Db 728 CAAAGAGGCTACCTATGTTGAACGTTGTGTAGCTTGTGTCAGTGGGAAGGCATTTCTT 787
 QY 198 eSerGlySerPheAlaAlaIlePheThrLeuLysLysArgGlyLeuMetProGlyLeuTh 218
 Db 788 TTCCGGTTCTTTTGGCTGATATTTGGCTCAAGAAACGAGGACTGATGCTCGCCCTCAC 847
 QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
 Db 848 ATTTTCTAATGAACCTTATTAGCAGAGATGAGGGTTTACACTGTGATTTTGTGCTGAT 907
 QY 238 tPheGlnTyrLeuValAenLysProSerGluGluArgValArgGluIleLeuValAspAl 258
 Db 908 GTTCAACACCTCGTACACAAACCATCGGAGGAGAGAGTAAGAGAAATAATTATCAATGC 967
 QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
 Db 968 TGTTCGATAGAACAGGAGTTCTCTACTGAGGCTTGCCTGTGAGAGCTCATTTGGGATGAA 1027
 QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuG1 298
 Db 1028 TTGCACCTTAATGAAGCAATACATTGAGTTTGTGGCAGACAGACTTATGCTGGAACTGGG 1087

QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG1 318
 DB 1088 TTTTACAGAGGTTTTCAGAGTAGAGAACCCATTGACTTATGGAGAAATTTCACTGGA 1147
 QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
 DB 1148 AGGAAAGACTAACTTCTTTGAGAGAGAGTAGCGGAGTATCAGAGATGGAGTGTATGTC 1207
 QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
 DB 1208 AAGTCAACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1247

RESULT 10

ID ADJ56536
 XX ADJ56536 standard; cDNA; 2216 BP.

AC ADJ56536;

XX 06-MAY-2004 (first entry)

DE Murine cDNA differentially expressed in MYCN activated cells SeqID 342.

XX mouse; murine; differential expression; transactivator; proto-oncogene;
 KW neuroblastoma; small cell lung cancer; cytostatic; gene therapy; ss;
 KW MYCN activated cell.

XX Mus musculus.

XX US2003119009-A1.

XX 26-JUN-2003.

XX 25-FEB-2002; 2002US-00084817.

XX 23-FEB-2001; 2001US-0270784P.

PA (STUA/) STUART S G.

PA (NUCH/) NUCHTERN J G.

PA (PLON/) PLON S E.

XX (SHOH/) SHOHET J M.

XX Stuart SG, Nuchtern JG, Plon SE, Shohet JW;

DR WPI; 2003-635698/60.

XX New genes regulated by MYCN activation, useful in gene therapy,
 PT particularly for treating a subject with e.g. neuroblastoma or
 PT cancers, or for diagnosing, staging or monitoring the treatment of the
 PT cancer.

XX Claim 1; SEQ ID NO 342; 27pp; English.

XX This invention relates to novel isolated cDNAs that are differentially
 CC expressed in MYCN activated cells. Specifically, it refers to
 CC polynucleotide sequences that exhibit differential expression patterns in
 CC cells activated by the transactivator MYCN, where MYCN is a proto-
 CC oncogene that is amplified in neuroblastoma cells and is common in small
 CC cell lung cancers. The present invention describes these cDNA molecules
 CC as useful for in hybridisation assays to detect expression of nucleic
 CC acids (or complementary nucleic acids) in a present in a given sample, as
 CC well as for screening assays by identifying molecules or compounds that
 CC specifically bind the cDNA as a ligand and modulate function or activity.
 CC Accordingly, these compositions exhibit cytostatic activity and can also
 CC be used for gene therapy purposes. This polynucleotide sequence is a cDNA
 CC that is differentially expressed in MYCN activated cells, given in an
 CC exemplification of the invention. NOTE: This sequence does not appear in
 CC the printed specification but has been obtained in electronic format from
 CC the US Patent Office at
 CC ftp.segdata.uspto.gov/sequence.html?DocID=20030119009.

XX Sequence 2216 BP; 593 A; 476 C; 524 G; 623 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1-42e-167 Length: 2216
 Score: 1460.50 Matches: 277
 Percent Similarity: 88.42% Conservative: 36
 Best Local Similarity: 78.25% Mismatches: 35
 Query Match: 80.20% Indels: 6
 DB: 10 Gaps: 1

US-10-698-228-1 (1-351) x ADJ56536 (1-2216)

QY 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
 DB 189 CCTGAGCGGAGCCCGGCTCCTGCCAGAACCGCGAGGAGATCTTCCAGGAGCCAC 248
 QY 23 nGluSerGluIleIysSer-----AsnGluGluProLeuLeuArgLysSe 38
 DB 249 GGAGCCGAAACTAAAGCAGCTGCCCGCGGTGGAGATGAGCCGCTGCTGAGAGAAA 308
 QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysG1 58
 DB 309 CCCCCCGCGCTTTGTTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 368
 QY 58 nAlaGlnAlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyr 78
 DB 369 GGCAGAGGCTTCTTTTGGACCGCGAGGAGGTGGACCTCTCCAAGGACATTTCAACATG 428
 QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAl 98
 DB 429 GGAATCCCTGAAACCCGAGGAGAGATATTTATATCCCATGTTCTTGGCTTTCTTTGAC 488
 QY 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
 DB 489 AAGCATGGATAGTAATAAGAACTTGGTGGACGATTTAGCCAGAGATTCAGATTAC 548
 QY 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
 DB 549 AGAAGCCCGCTGTTCTATGGCTTCCAAATTCGCAATGCGAAACATACATCTCTGAATGTA 608
 QY 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
 DB 609 TAGTCTCTTTATTGACACTTACATAAAAGATCCCAAGAAAGAGGAATTTCTTCAATGC 668
 QY 158 aIleGluThrMetProTyrValLysLysLysAlaAspTyrAlaLeuArgTyrPheAlaAs 178
 DB 669 CATTGAAACCATGCTGCTGTCAGAGAGAGGACAGACTGGGCTTGCCTGGATGGGA 728
 QY 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePh 198
 DB 729 CAAAGAGGCTACCTATGCTGAAACGTTGTAGCTTTTGTCTGCTGAGTGGAGGCAATTTCTT 788
 QY 198 eSerGlySerPheAlaAlaIlePheThrLeuLysLysArgGlyLeuMetProGlyLeuTh 218
 DB 789 TTCGGTTCCTTTTTCGGTGCATATTTCTGGCTCAAGAAACGAGGACTGATGCTCGGCTCAC 848
 QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
 DB 849 ATTTTCTAATGAATTTATAGCAGATAGAGGTTTACACTGTGATTTTGTGCTGCTGAT 908
 QY 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluLileIleValAspAl 258
 DB 909 GTTCAACACCTGGTACACAAACCATCGAGGAGAGAGTAAGAGAAATAATATCAATGC 968
 QY 258 aValLysIleGluGlnGluPheThrGluAlaLeuProValGlyLeuIleGlyMeAs 278
 DB 969 TGTTCCGATAGAACAGGAGTTCTCTCACTGAGGCTTGCCTGTGAGGCTCATTTGGATGAA 1028
 QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuG1 298
 DB 1029 TTGCACCTCTAATGAAGCAATACATTTGATTTTGTGGCAGACAGACTTATGCTGGAACTGG 1088
 QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG1 318
 DB 1089 TTTTAGCAAGGTTTTCAGAGTAGAGAACCCATTGACTTTATGAGAGATATTTCACTGGA 1148

QY 318 uGlyLysThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
 DB 1149 AGGAAGACTAACTCTTTGAGACAGAGTAGGCGAGTATCAGAGGATGGGAGTGATGTC 1208
 QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
 DB 1209 AAGTCCAACAGAGAAATCTTTTACCTTGATGCTGACTTC 1248

RESULT 11

RAAS44917/c
 ID AAS44917 standard; DNA; 2482 BP.
 AC AAS44917;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Human contig polynucleotide sequence #170.
 XX
 KW Mammal; human; rhesus monkey; baker's yeast; fission yeast; Norway rat;
 KW mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;
 KW cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;
 KW nervous system disorder; inflammatory disorder; cell differentiation; ds;
 KW angiogenesis; stem cell growth factor; activin; inhibin; cartilage; burn;
 KW genetic disorder; bone regeneration; tendon; ligament; tissue repair;
 KW cytostatic; antirheumatic; antiarthritic; vulnary; antiinflammatory;
 KW antibacterial; immunosuppressive; vasotropic; antiparkinsonian;
 KW neuroprotective; osteopathic; antidiabetic; antiasthmatic; antiallergic;
 KW immunostimulant; analgesic; gene therapy.
 XX

OS Homo sapiens.
 OS Synthetic.

XX WO200164834-A2.

XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US004926.

XX 28-FEB-2000; 2000US-00515126.

XX 18-MAY-2000; 2000US-00577409.

XX 17-JUN-2000; 2000US-00597707.

XX 14-JUL-2000; 2000US-00616807.

XX 19-SEP-2000; 2000US-00664641.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QX, Ren F;
 PI Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;
 PI Drmanac R;

XX WPI; 2001-589862/66.
 DR P-PSDB; AAU28017.

XX Novel polypeptides and nucleic acids obtained from cDNA libraries
 PT prepared from various human tissues, for diagnosis, treatment of cancer,
 PT neurological, inflammatory disorders and for use in arrays for detection.

XX Claim 1; SEQ ID NO 514; 153pp; English.

XX Sequences AAS44576-AAS44919 represent full-length polynucleotides and
 CC contig polynucleotides encoding polypeptides of the invention. The DNA
 CC and protein sequences are useful for the treatment, diagnosis and
 CC prevention of various types of disorder in a mammalian subject such as a
 CC human, dog, monkey, mouse, hamster or rat. The disorders include cancers
 CC such as leukaemia, lymphoma and neuroblastoma, autoimmune disorders such
 CC as multiple sclerosis, connective tissue disease, rheumatoid arthritis,
 CC diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system
 CC disorders such as Parkinson's disease, Alzheimer's disease, Huntington's
 CC chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and
 CC Werner's disease, inflammatory disorders such as nephritis, Crohn's
 CC disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory
 CC bowel disease. The sequences exhibit activity relating to angiogenesis,

CC cell proliferation, cell differentiation, stem cell growth factor,
 CC activin or inhibin. Therefore, they can be used to manipulate stem cells
 CC in culture to give rise to neuroepithelial cells that can be used to
 CC augment or replace cells damaged by illness, accidental damage or genetic
 CC disorders. The sequences may also be used for regeneration of bone,
 CC cartilage, tendons and ligaments and in tissue repair and burn healing.
 CC Note: Some sequences for this patent did not form part of the printed
 CC specification, but were obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences

XX
 SQ Sequence 2482 BP; 707 A; 598 C; 522 G; 655 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.68e-167 Length: 2482
 Score: 1460.50 Matches: 277
 Percent Similarity: 88.42% Conservative: 36
 Best Local Similarity: 78.25% Mismatches: 35
 Query Match: 80.20% Indels: 6
 Gaps: 1

US-10-698-228-1 (1-351) x AAS44917 (1-2482)

QY 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerAspThrAs 23
 DB 2181 CCTGAGCGGACCCCGCGTCTCTGGCCAGCAAGACCGCGAGGAGGATCTTCCAGGAGCCAC 2122
 QY 23 nGluSerGluLeuLysSer-----AsnGluGluProLeuLeuArgLysSe 38
 DB 2121 GGAGCGCAAACTAAAGCAGCTGCCCGCGCGTGGAGGATGAGCGCTGCTGAGAGAAA 2062
 QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysG1 58
 DB 2061 CCCCCCGCGCTTGTCTATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAGAA 2002
 QY 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
 DB 2001 GGCAGAGGCTCTCTTTTGGACCGCGCGAGGAGGTTGACCTCTCCAAGGACATTCAGCCTG 1942
 QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAl 98
 DB 1941 GGAATCCCTGAAACCCGAGGAGAGATATTTTATATCCATGTTCTGGCTTTCTTTGACG 1882
 QY 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
 DB 1881 AAGCGATGGCATAGTAAATGAAACCTTGGTGGAGCGGATTTAGCCCAAGATTCAGATTAC 1822
 QY 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
 DB 1821 AGRAGCCCGCTGTTCTATGGCTTCCAAATTTGCCATGGAACACATACATCTGGAATGTA 1762
 QY 138 rSerLeuLeuLeuAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
 DB 1761 TAGTCTTCTATTGACACTTACATAAAAGATCCCAAAAGAAAGGAATTTCTTCTCAATGC 1702
 QY 158 alLeGluThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
 DB 1701 CATTGAAACGATGCTCTTGTGTCAGAGAGGAGGAGCTGGGCTTGGCTGATGGGGA 1642
 QY 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePh 198
 DB 1641 CAAGAGGCTACTATGTTGTAACGTTGTTAGGCTTTGCTGCGATGGAGGCAATTTCTT 1582
 QY 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
 DB 1581 TTCCGGTCTTTTTCGCTCATATTCTGGCTCAAGAAACGAGGAGCTGATGCTGCGCTCAC 1522
 QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
 DB 1521 ATTTTCTAATGAATTTATTAGCAGATGAGGGTTTACACTGTGATTTTGTCTCCCTGAT 1462
 QY 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
 DB 1461 GTTCAACACCTGGTATACACAACCATCGGAGGAGAGATTAAGAGAAATATTATCAATGC 1402

```

QY 258 avallysileGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
DB 1401 TGTTCGATAGACAGGAGTCTCTACTAGAGCCTTGCTGTGAAGCTCATTTGGATGAA 1342
QY 278 nCysIleLeuMetLysGlnTyRileGluPheValAlaAspArgLeuValGluLeuGl 298
DB 1341 TTGCACTCTAATGAAGCAATACATTGAGTTTGTGGCAGACAGACTTATGCTGGAACGGG 1282
QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGl 318
DB 1281 TTTTAGCAGAGGTTTTCAGAGTAGAGAACCAATTGACTTTATGGAGAAATATTTCACCTGA 1222
QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyRValGlnArgPheAlaValMetAl 338
DB 1221 AGGAAGACTAACTTCTTTGAGAAGAGATAGGCAGATATCAGAGATGGGAGTGAATGTC 1162
QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB 1161 AAGTCCAAACAGAGAATTCTTTTACCTTGGATGCTGACTTC 1122

RESULT 12
ABL65414
ID ABL65414 standard; DNA; 2500 BP.
XX
AC ABL65414;
XX
DT 15-MAY-2002 (first entry)
XX
DE Lung cancer related gene sequence SEQ ID NO:3751.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US010838.
XX
PR 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0233133P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.

```

```

PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
XX (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
DR WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX
PS Claim 1; SEQ ID NO 3751; 4pp; English.
XX
CC The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms'
CC tumour
XX
SQ Sequence 2500 BP; 679 A; 522 C; 595 G; 704 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,7e-167 Length: 2500
Score: 1460.50 Matches: 277
Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
DB: Gaps: 1

US-10-698-228-1 (1-351) x ABL65414 (1-2500)
QY 4 ProGluArgProGluAlaAlaGlyLeuAspGluArgSerSerSerAspThrAs 23
DB 302 CCTGAGCGGAGCCCGGCTCTGCGCCAGACGCGGAGGAGGATCTTCCAGAGGCCAC 361
QY 23 ngluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
DB 362 GGAGCCGAAACTAAAGCAGACTGCCCGCGGAGGATGAGCGGCTGCTGAGAGAAA 421
QY 38 rSerArgArgPheValIlePheProIleGlnTyRProAspIleTrpLysMetTrpLysGl 58
DB 422 CCCCCCGCGCTTGTCTATCTTCCCATCGAGTACCAGATATCTGGCAGATGATATAGAA 481
QY 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAlaAspLeuSerLysAspLeuProHisTr 78
DB 482 GGCAGAGGCTTCTCTTTTGGACCCCGGAGGAGGTTGACCTCTCCAAAGGACATTCAGCACTG 541

```

QY 78 pAenLysLeuLysAlaAspGluLysTyrPheLeSerHisIleLeuAlaPhePheAlaAl 98
Db 542 GGAATCCCTGAAACCGAGAGAGATATTTATATCCATGTTCTGGCTTTCTTCGAGC 601
QY 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGluGluValGlnValPr 118
Db 602 AAGCGATGCGATAGTAAATGAAACATTTGGTGAGCGATTTAGCCAAAGAGTTCCAGATTAC 661
QY 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
Db 662 AGAAGCCCGCTGTTCTATGGCTTCGAATTCGATGGAACATACATCTTCGAATGTA 721
QY 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
Db 722 TAGTCTTCTTATTGACACTTACATAAAGATCCCAAGAAAGGGAATTTCTCTTCAATGC 781
QY 158 alleGluThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
Db 782 CATTGAACGATGCTTGTGTCAAGAAAGAGGAGAGCTGGGCTTCGCTGGATTGGGA 841
QY 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePh 198
Db 842 CAAAGAGGCTACTGTGTGGAACGCTTTGTAGCTTTGTCTGAGTGAAGGCAATTTCTT 901
QY 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
Db 902 TTCGGTCTTTTGGCTCGATATTTCTGGCTCAAGAAACGAGGACTGATGCTGGCTCAC 961
QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
Db 962 ATTTCTTAATGAATATTATGACAGATGAGGGTTTACACTGTGTATTTTGTCTGCTGAT 1021
QY 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
Db 1022 GTTCAACACCTGGTACACAAACCATCGAGAGAGAGTAAAGAGAAATAATTATCAATGC 1081
QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
Db 1082 TGTTGGATAGAACAGGAGTTCTACTAGGCTTGGCTGTGAAGCTCATTGGGATGA 1141
QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGl 298
Db 1142 TTGCACTTAATGAAGCAATACATTGAGTTTGTGGCAGACAGACTTATGCTGGAATGGG 1201
QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGl 318
Db 1202 TTTTAGCAAGGTTTTTCAGAGTAGAAGAACCCATTTCAGCTTTATGGAGAAATATTTCACTGGA 1261
QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
Db 1262 AGAAGAGCTAACTTTTGAAGAGAGATGGGAGTATCAGAGATGGGAGTATGATGTC 1321
QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1322 AAGTCCACAGAGATTTCTTTACCTTGGATGCTGACTTC 1361

RESULT 13

ABL66517

ID ABL66517 standard; DNA; 2500 BP.

XX AC ABL66517;

XX AC ABL66517;

XX DT 15-MAY-2002 (first entry)

XX DE Lung cancer related gene sequence SEQ ID NO:4854.

XX KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;

XX KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;

XX KW cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;

XX KW gene; db.

XX OS Homo sapiens.

CC The present invention describes a method (M1) for screening for an anti-neoplastic agent. The method involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, determining a change in expression of at least one gene (I) of a signature gene set, where (I) comprises a sequence (S) selected from 8417 sequences (given in ABL61664 to ABL70110), or is at least 95% identical to (S), where a change in expression is indicative of anti-neoplastic activity. (I) has cytostatic activity and can be used in gene therapy. M1 can be used for screening an anti-neoplastic agent, and can be used for producing a product which is the data collected with respect to the anti-neoplastic agent as a result of M1, and the data is sufficient to convey the chemical structure and/or

WO200194629-A2.

13-DEC-2001.

30-MAY-2001; 2001WO-US010838.

05-JUN-2000; 2000US-0209473P.

05-JUN-2000; 2000US-0209531P.

PR 18-SEP-2000; 2000US-0231133P.

PR 18-SEP-2000; 2000US-0233617P.

PR 20-SEP-2000; 2000US-0234009P.

PR 20-SEP-2000; 2000US-0234034P.

PR 20-SEP-2000; 2000US-0234052P.

PR 22-SEP-2000; 2000US-0234509P.

PR 22-SEP-2000; 2000US-0234567P.

PR 25-SEP-2000; 2000US-0234923P.

PR 25-SEP-2000; 2000US-0234924P.

PR 25-SEP-2000; 2000US-0235077P.

PR 25-SEP-2000; 2000US-0235082P.

PR 25-SEP-2000; 2000US-0235134P.

PR 25-SEP-2000; 2000US-0235280P.

PR 26-SEP-2000; 2000US-0235637P.

PR 26-SEP-2000; 2000US-0235638P.

PR 27-SEP-2000; 2000US-0235711P.

PR 27-SEP-2000; 2000US-0235720P.

PR 27-SEP-2000; 2000US-0235840P.

PR 27-SEP-2000; 2000US-0235863P.

PR 28-SEP-2000; 2000US-0236028P.

PR 28-SEP-2000; 2000US-0236032P.

PR 28-SEP-2000; 2000US-0236033P.

PR 28-SEP-2000; 2000US-0236034P.

PR 28-SEP-2000; 2000US-0236109P.

PR 28-SEP-2000; 2000US-0236111P.

PR 29-SEP-2000; 2000US-0236842P.

PR 29-SEP-2000; 2000US-0236891P.

PR 02-OCT-2000; 2000US-0237172P.

PR 02-OCT-2000; 2000US-0237173P.

PR 02-OCT-2000; 2000US-0237278P.

PR 02-OCT-2000; 2000US-0237294P.

PR 02-OCT-2000; 2000US-0237295P.

PR 03-OCT-2000; 2000US-0237316P.

PR 03-OCT-2000; 2000US-0237425P.

PR 03-OCT-2000; 2000US-0237598P.

PR 03-OCT-2000; 2000US-0237604P.

PR 03-OCT-2000; 2000US-0237606P.

PR 03-OCT-2000; 2000US-0237608P.

PR 01-NOV-2000; 2000US-0244867P.

PR 01-NOV-2000; 2000US-0245084P.

(AVAL-) AVALON PHARM.

Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;

Soppet DR, Weaver Z;

WPI; 2002-188264/24.

Screening for anti-neoplastic agent involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, and determining a change in expression of a gene of a signature gene set.

Claim 1; SEQ ID NO 4854; 44pp; English.

The present invention describes a method (M1) for screening for an anti-neoplastic agent. The method involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, determining a change in expression of at least one gene (I) of a signature gene set, where (I) comprises a sequence (S) selected from 8417 sequences (given in ABL61664 to ABL70110), or is at least 95% identical to (S), where a change in expression is indicative of anti-neoplastic activity. (I) has cytostatic activity and can be used in gene therapy. M1 can be used for screening an anti-neoplastic agent, and can be used for producing a product which is the data collected with respect to the anti-neoplastic agent as a result of M1, and the data is sufficient to convey the chemical structure and/or

CC properties of the agent. M1 can be used in the treatment of cancer such
 CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
 CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
 CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
 CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
 CC tumour

SQ Sequence 2500 BP; 679 A; 522 C; 595 G; 704 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1,7e-167 Length: 2500
 Score: 1460.50 Matches: 277
 Percent Similarity: 88.42% Conservative: 36
 Best Local Similarity: 78.25% Mismatches: 35
 Query Match: 80.20% Indels: 6
 DB: 6 Gaps: 1

US-10-698-228-1 (1-351) x ABL6517 (1-2500)

QY	4	ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs	23
DB	302	CCTGAGCGGACCCGGCTCTGGCCAGCAAGACCGGAGGAGGATCTTCAGGAGCCAC	361
QY	23	nGluSerGluLeuLysSer-----AenGluGluProLeuLeuArgLysSe	38
DB	362	GGAGCCGAAACTAAAGCAGCTGCCCGCGGTGGAGGATGAGCGCTCTCGAGAGAAA	421
QY	38	rSerArgArgPheValIlePheProIleGlnTyrProAspIleTyrPheMetTyrLysG	58
DB	422	CCCCCGCGCTTGTTCATCTCCCATCGAGTACCATGATATCTGCAGATGTATAAGAA	481
QY	58	nAlaGlnAlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTr	78
DB	482	GGCAGAGGCTTCCTTTGGACCGCCGAGGAGGTGTACCTCTCCAGGACATTCAGCACG	541
QY	78	pAsnLysLeuLysAlaAspGluLysTyrPheLeSerHisIleLeuAlaPhePheAlaAl	98
DB	542	GGAAATCCCTGAAACCGGAGGAGATATTTATATCCCATGTTCTGGCTTCTTTCGAGC	601
QY	98	sSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr	118
DB	602	AAGCGATGTCATAGTAAATGAAACCTTGGTGAGCGGATTTAGCCAAAGAGTTTCAGAT	661
QY	118	oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy	138
DB	662	AGAAGCCGCTGTTTCTATGGCTTCCAAATGCCATGGAAACATACATCTGAAATGA	721
QY	138	rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPhePheAsnAl	158
DB	722	TAGTCTCTTATGTACACTTACATAAAGATCCCAAGAAAGGGAATTTCTTCAATGC	781
QY	158	alleGluThrMetProTyrValLysLysLysAlaAspTyrAlaLeuArgTyrIleAlaAs	178
DB	782	CATTGAAACGATGCTTGTGTCAAGAAGAGGCAGACTGGGCTTCGCTGGATGGGGA	841
QY	178	pArgLysSerThrPheGlyGluArgValValAlaPheAlaValGluGluValPhePh	198
DB	842	CAAGAGGCTACTATGGTGAACGTTGTAGCTTGTGCTGAGTGAAGGCAATTTCTTT	901
QY	198	sSerGlySerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuTh	218
DB	902	TTCCGGTCTTTTGGTGCATATTTCTGGCTCAAGAAACGAGGACTGATCGCTGGCTCAC	961
QY	218	rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe	238
DB	962	ATTTTCTAATGAATATTAGCAGAGATGAGGTTTACACTGTGATTTTGTCTTCCTGAT	1021
QY	238	tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl	258
DB	1022	GTTCAACACCTGGTACAAACCATTCGAGGAGGAGTAAAGAGAAATAATATCAATGC	1081
QY	258	aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs	278

DB	1082	TGTTGGATAGAACAGAGGTTCTCTCACTAGCGGCTTGCCTGTGTGAAGCTCATTTGGATGAA	1141
QY	278	nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuG	298
DB	1142	TTGCATCTTAATGAACATACATTTAGTTTGTGGCAGACAGACTTATGCTTGGAACTGGG	1201
QY	298	yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG	318
DB	1202	TTTTAGCAAGTTTTCAGAGTAGAGAACCCATTTGACTTTATCGAGAAATATTTCACTGA	1261
QY	318	uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl	338
DB	1262	AGGAAAGACTAATCTCTTTTGAGAAGAGAGTAGGCGAGTATCAGAGGATGGGAGTGATGC	1321
QY	338	aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe	351
DB	1322	AGTCCCAACAGAAATCTTTTACCTTGGATGCTGACTTC	1361

RESULT 14
 ABL65859
 ID ABL65859 standard; DNA; 2500 BP.
 XX
 AC ABL65859;
 DX
 DT 15-MAY-2002 (first entry)
 XX
 DE Lung cancer related gene sequence SEQ ID NO:4196.
 XX
 KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200194629-A2.
 XX
 PD 13-DEC-2001.
 XX
 PF 30-MAY-2001; 2001WO-US010838.
 PR
 PR 05-JUN-2000; 2000US-0209473P.
 PR 05-JUN-2000; 2000US-0209531P.
 PR 18-SEP-2000; 2000US-0233133P.
 PR 18-SEP-2000; 2000US-0233617P.
 PR 20-SEP-2000; 2000US-0234009P.
 PR 20-SEP-2000; 2000US-0234034P.
 PR 20-SEP-2000; 2000US-0234052P.
 PR 22-SEP-2000; 2000US-0234509P.
 PR 22-SEP-2000; 2000US-0234567P.
 PR 25-SEP-2000; 2000US-0234923P.
 PR 25-SEP-2000; 2000US-0234924P.
 PR 25-SEP-2000; 2000US-0235077P.
 PR 25-SEP-2000; 2000US-0235082P.
 PR 25-SEP-2000; 2000US-0235134P.
 PR 25-SEP-2000; 2000US-0235280P.
 PR 26-SEP-2000; 2000US-0235637P.
 PR 26-SEP-2000; 2000US-0235638P.
 PR 27-SEP-2000; 2000US-0235711P.
 PR 27-SEP-2000; 2000US-0235720P.
 PR 27-SEP-2000; 2000US-0235840P.
 PR 27-SEP-2000; 2000US-0235863P.
 PR 28-SEP-2000; 2000US-0236028P.
 PR 28-SEP-2000; 2000US-0236032P.
 PR 28-SEP-2000; 2000US-0236033P.
 PR 28-SEP-2000; 2000US-0236034P.
 PR 28-SEP-2000; 2000US-0236109P.
 PR 28-SEP-2000; 2000US-0236111P.
 PR 29-SEP-2000; 2000US-0236842P.
 PR 29-SEP-2000; 2000US-0236891P.
 PR 02-OCT-2000; 2000US-0237172P.
 PR 02-OCT-2000; 2000US-0237173P.
 PR 02-OCT-2000; 2000US-0237278P.

PR 02-OCT-2000; 2000US-0237294P.
 PR 02-OCT-2000; 2000US-0237295P.
 PR 02-OCT-2000; 2000US-0237316P.
 PR 03-OCT-2000; 2000US-0237425P.
 PR 03-OCT-2000; 2000US-0237598P.
 PR 03-OCT-2000; 2000US-0237604P.
 PR 03-OCT-2000; 2000US-0237606P.
 PR 03-OCT-2000; 2000US-0237608P.
 PR 01-NOV-2000; 2000US-0244867P.
 PR 01-NOV-2000; 2000US-0245084P.
 XX (AVAL-) AVALON PHARM.
 PA Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrihan S;
 PI Soppet DR, Weaver Z;
 XX WPI; 2002-188264/24.
 DR
 XX Screening for anti-neoplastic agent involves exposing cells to a chemical
 PT agent to be tested for anti-neoplastic activity, and determining a change
 PT in expression of a gene of a signature gene set.
 XX
 PS Claim 1; SEQ ID NO 4196; 44pp; English.
 CC The present invention describes a method (M1) for screening for an anti-
 CC neoplastic agent. The method involves exposing cells to a chemical agent
 CC to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening an
 CC anti-neoplastic agent, and can be used for producing a product which is
 CC the data collected with respect to the anti-neoplastic agent as a result
 CC of M1, and the data is sufficient to convey the chemical structure and/or
 CC properties of the agent. M1 can be used in the treatment of cancer: such
 CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
 CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
 CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
 CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
 CC tumour
 XX
 SQ Sequence 2500 BP; 679 A; 522 C; 595 G; 704 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.7e-167 Length: 2500
 Score: 1460.50 Matches: 277
 Percent Similarity: 88.42% Conservative: 36
 Best Local Similarity: 78.23% Mismatches: 35
 Query Match: 80.20% Indels: 6
 DB: 6 Gaps: 1
 US-10-698-228-1 (1-351) x ABL65859 (1-2500)
 QY 4 ProGluArgProGluAlaAlaGlyLeuAspGluArgSerSerAspThrAs 23
 DB 302 CCGTGGCGGAGCCCGCGTCTGGCGAGCAAGACCGCGAGGAGGATCTTCAGAGGCCAC 361
 QY 23 nGluSerGluIleLysSer-----AanGluGluProLeuLeuArgLysSe 38
 DB 362 GGAGCGGAAACTAAGACAGTGCCTGGCGGAGGAGGAGCGCTGCTGAGAGAA 421
 QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysG1 58
 DB 422 CCCCAGCGCGCTTGTTCATCTCCCATCGAGTACCATGATATCTGCAGATGTATAAGAA 481
 QY 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
 DB 482 GCGAGAGGCTTCCTTTTGGACCGCGGAGGAGGTTGACCTCTCCAGAGCAATTCAGCACTG 541
 QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAl 98
 DB 542 GGAATCCCTGAAACCCGAGGAGAGATATTTATATCCCAATGTTCTGGCTTTCTTTCAGC 601

QY 98 aserAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
 DB 602 AAGCGATGGCATAGTAATAAGAAACTTGGTGGAGCGATTAGCCAAAGATTTCAGATTAC 661
 QY 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
 DB 662 AGAAGCCCGCTGTTTCTATGGCTTCCAAATTCGCAATGCGAAACATACATCTCTGAATGTA 721
 QY 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysIlysbArgGluPheLeuPheAsnAl 158
 DB 722 TAGTCTCTTATTGACACTTACATAAAGATCCCAAGAAAGGGAATTTCTTCAATGC 781
 QY 158 alIeGluThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
 DB 782 CATTAACACGATGCTTGTGTCAAGAAAGAGGAGACTGGGCTTGGCTGGAATGGGGA 841
 QY 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePh 198
 DB 842 CAAAGAGGCTACCTATGTTGTAACGTTGTAGCCTTTGCTGCACTGGAGGCAATTTCTT 901
 QY 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
 DB 902 TTCGGCTTCTTTTGGCTCGATATTCTGGCTCAAGAAACGAGGACTGATGCTGGCTCAC 961
 QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
 DB 962 ATTTTCTTAATGAATATTATTAGCAGAGATGAGGGTTTACACTGTGATTGTGCTGCTGAT 1021
 QY 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
 DB 1022 GTTCAACACCTGTGTACACAAACCATCGGAGGAGAGTAGAGAAATAATATCAATGC 1081
 QY 258 aValLysIleGluGlnGlnPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
 DB 1082 TGTTCGATAGAACAGGAGTTCTCTCACTGAGGCCCTTCTGCTGTAAGCTCATTTGGATGAA 1141
 QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuG1 298
 DB 1142 TTGCACCTCTAATGAAGCAATACATATTGATTGTGTGGCAGACAGACTTATGCTGGAACCTGG 1201
 QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG1 318
 DB 1202 TTTTAGCAAGGTTTTCAGATAGAACCCCATTTGACTTTATGGAGAAATATTTTCACTGGA 1261
 QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
 DB 1262 AGGAAGACTAACTTCTTTGAGAAGAGAGTAGGCGAGTATCAGAGGATGGGAGTGATGTC 1321
 QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
 DB 1322 AAGTCCACAGAGAAATTTCTTTTACCTTGGATGCTGACTTC 1361
 RESULT 15
 ABX10335
 ID ABX10335 standard; DNA; 2500 BP.
 XX
 AC ABX10335;
 XX
 DT 28-JAN-2003 (first entry)
 XX
 DE DNA encoding protein differentially regulated in prostate cancer #4.
 KW Prostate cancer; gene expression; differential regulation;
 KW molecular marker; drug target; cancer detection; cancer diagnosis;
 KW cancer staging; cancer grading; cancer assessing; cancer monitoring;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200281638-A2.
 XX
 PD 17-OCT-2002.

XX 08-APR-2002; 2002WO-US010824.
 XX
 PR 06-APR-2001; 2001US-0281731P.
 PR 06-APR-2001; 2001US-0281732P.
 XX (ORIG-) ORIGENE TECHNOLOGIES INC.
 XX
 XX Sun Z, Jay G;
 DR WPI; 2003-058520/05.
 DR P-PSDB; ABU07433.
 XX
 PT Novel genes which are differentially regulated in prostate cancer, useful
 PT for diagnosing prostate cancer in prostate tissue sample and assessing
 PT therapeutic or preventive intervention in prostate cancer patients.
 XX
 PS Claim 1; Page 71-72; 416pp; English.
 CC The invention describes genes (I) which are differentially regulated in
 CC prostate cancer. (I) is useful for diagnosing a prostate cancer in a
 CC sample comprising prostate tissue, which involves determining the number
 CC of target genes which are differentially-regulated in the sample, where
 CC the number is indicative of the probability that the sample comprises
 CC prostate cancer. (I) is useful for assessing a therapeutic or preventive
 CC intervention in a subject having a prostate cancer, which involves
 CC determining the expression levels in a sample comprising prostate tissue
 CC of target genes which are differentially-regulated in prostate cancer.
 CC Preferably, the expression levels of at least 10 genes are determined.
 CC (I) is also useful for identifying agents that modulate a biological
 CC activity of a polypeptide differentially-regulated in prostate cancer
 CC cells, which involves contacting a polypeptide differentially-regulated
 CC in prostate cancer cells with a test agent under conditions effective for
 CC the test agent to modulate a biological activity of the polypeptide, and
 CC determining whether the test agent modulates the biological activity. (I)
 CC is useful as molecular markers, as drug targets, and for detecting,
 CC diagnosing, staging, grading, assessing, monitoring, prognosticating,
 CC preventing or treating, determining predisposition to diseases and
 CC conditions especially relating to prostate cancer. (I) and its expression
 CC products are used in the diagnostic test to assay for presence of cancer
 CC e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in
 CC blood etc. (I) is useful for assessing cancer e.g., to determine the type
 CC of cancer, its stage of development, the nature of genetic defect, etc.
 CC The polypeptide encoded by (I) can be used as target for therapy or drug
 CC discovery. (I) can also be used for expressing the polypeptide and thus
 CC for searching specific binding partners of the polypeptide. (I) is useful
 CC in therapeutic applications to treat prostate cancer. The identification
 CC of specific genes, and groups of genes, expressed in pathways
 CC physiologically relevant to prostate cancer permits the definition of
 CC functional and disease pathways and the delineation of targets in these
 CC pathways which are useful in diagnostic, therapeutic, and clinical
 CC applications. This sequence encodes a protein differentially regulated in
 CC prostate cancer
 XX
 SQ Sequence 2500 BP; 679 A; 522 C; 595 G; 704 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1.7e-167 Length: 2500
 Score: 1460.50 Matches: 277
 Percent Similarity: 88.42% Conservative: 36
 Best Local Similarity: 78.25% Mismatches: 35
 Query Match: 80.20% Indels: 6
 DB: 8 Gaps: 1

US-10-698-228-1 (1-351) x ABX10335 (1-2500)

QY 4 ProGluArgProGluAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
 DB 302 CTGAGCGGAGCCCGGCTCTGCGCCAGCAAGACCGGAGGAGATCTTCAGGAGCCAC 361
 QY 23 nGluSerGluLeuLysSer-----AsnGluGluProLeuLeuArgLysSe 38
 DB 362 GGAGCGGAAACTAAAGCAGCTGCCCGCGCTGGAGGATGAGCCGCTCTGAGAGAAA 421

QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysG1 58
 DB 422 CCCCCCGCGCTTGTCTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGATATAGAA 481
 QY 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLysSerLysLeuProHisTr 78
 DB 482 GGCAGAGGCTTCTCTTTGGACCGCGAGGAGGTGACCTCTCCAAGGACATTCAGCATG 541
 QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAl 98
 DB 542 GGAATCCCTGAAACCCGAGGAGAGATATTTATATCCATGTTCTTGGCTTCTTTGACG 601
 QY 98 aSerAspGlyIleValAsnGlnAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
 DB 602 AAGCGATGGCATAGTAATAAGAACTTGGTGGAGGATTTAGCCAAAGATTCAGATTAC 661
 QY 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
 DB 662 AGAAGCCCGCTGTTTCTATGGCTTCCAAATTCGCATGGAACATACATCTGAAATGTA 721
 QY 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
 DB 722 TAGTCTTCTTATTGACACTTACATAAAGATCCCAAGAAAGGGAATTTCTTTCATGTC 781
 QY 158 aIleGluThrMetProTyrValLysLysAlaAspTyrAlaLeuArgTyrIleAlaAs 178
 DB 782 CATTGAACCATGCTTGTCTCAAGAAAGAGGAGACTGGGCTTGGCTGGATGGGGA 841
 QY 178 pArgLysSerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePh 198
 DB 842 CAAAGAGGCTACCTGTTGCTGCAACGTTGTAGCTTTGCTGCTGCAAGGCAATTTCTT 901
 QY 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
 DB 902 TTCGGTCTCTTTTGGCTCGATATTTCTGGCTCAAGAAACGAGGACTGATGCTGCCCTCAC 961
 QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
 DB 962 ATTTCTAATGAACCTTATTAGCAGAGATGAGGTTTACCTGTGATTTGCTTGCCTGAT 1021
 QY 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
 DB 1022 GTTCAACACCTGTCACAAACCATCGGAGGAGAGATGAAGAAATAATTATCAATGC 1081
 QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
 DB 1082 TGTTCGGATAGAACAGGAGTTCTCTCATGAGGCTTGCCTGTGAAGCTCATTTGGATGAA 1141
 QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluG1 298
 DB 1142 TTCACCTCTAATGAAGCAATACATTGATTTGTGGCAGACAGACTTATGCTGGAACTGGG 1201
 QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG1 318
 DB 1202 TTTTAGCAGGTTTTCAGATGAGAACCCATTGTACTTTATGGAGATAATTTTCACTGA 1261
 QY 318 uGlyLysThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
 DB 1262 AGGAAAGACTAACTCTTTTGAAGAGAGATGAGCGAGTATCAGAGGATGGGAGTGATGTC 1321
 QY 338 aGluThrThrAspAsnValPheThrIleuAspAlaAspPhe 351
 DB 1322 AGTCCCAACAGAGAAATCTTTTACCTTGGATGTGACTTC 1361

Search completed: October 30, 2005, 07:11:59
 Job time : 466 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 30, 2005, 07:01:57 ; Search time 2584 Seconds

(without alignments)
5170.495 Million cell updates/sec

Title: US-10-698-228-1

Perfect score: 1821

Sequence: 1 MGDPERPEAGLQDQERSSS.....QRFVMAETTDNVFLDADF 351

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q/cgn2.1/USPT001/US10698228/runat 26102005 100715 5117/app_query.fasta_1.519
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10698228@cgn.1.1.2607 @runat 26102005 100715 5117 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1821	100.0	3292	3	CR617553 full-leng
2	1821	100.0	4748	3	HSW802033
3	1791	98.4	4650	3	BC042948 Homo sapi
4	1737.5	95.4	3501	3	CR627376 Homo sapi
5	1591	87.4	1083	1	AL547501
6	1585	87.0	1056	9	AY398973 Homo sapi
7	1468	80.6	1043	9	AY398975 Mus muscu
8	1460.5	80.2	1571	3	CR603461 full-leng
9	1460.5	80.2	1573	3	CR625489 full-leng

10	1460.5	80.2	1582	3	CR608076	full-leng
11	1460.5	80.2	1588	3	CR602054	full-leng
12	1460.5	80.2	1592	3	CR621427	full-leng
13	1460.5	80.2	1600	3	CR604378	full-leng
14	1460.5	80.2	1605	3	CR590959	full-leng
15	1460.5	80.2	1612	3	CR614990	full-leng
16	1460.5	80.2	1613	3	CR609838	full-leng
17	1460.5	80.2	1623	3	CR618451	full-leng
18	1460.5	80.2	1630	3	CR602150	full-leng
19	1460.5	80.2	1796	3	CR596700	full-leng
20	1443	79.2	2088	3	AK088907	Mus muscu
21	1384	76.0	879	5	BQ441857	AGENCOURT
22	1354	74.4	892	7	CN163214	952695 MA
23	1328.5	73.0	997	4	BM468712	AGENCOURT
24	1316	72.3	1008	9	AY398974	Pan trogl
25	1277	70.1	932	5	BUI96941	AGENCOURT
26	1275	70.0	914	7	CF995079	AGENCOURT
27	1257.5	69.1	946	7	CN024619	AGENCOURT
28	1235	67.8	870	7	CO648380	ILLUMIGEN
29	1222	67.1	836	7	CR416843	CR416843
30	1221	67.1	893	5	BUI90680	AGENCOURT
31	1209.5	66.4	923	7	CR580780	CR580780
32	1209	66.4	1076	4	BM460735	AGENCOURT
33	1203	66.1	801	5	BP680185	BP680185
34	1197.5	65.8	933	6	CA983633	AGENCOURT
35	1195	65.6	1111	5	BM914217	AGENCOURT
36	1189	65.3	909	6	CA981614	AGENCOURT
37	1188	65.2	804	6	CD656661	AGENCOURT
38	1186.5	65.2	1028	5	BQ050629	AGENCOURT
39	1186	65.1	879	5	BUI89053	AGENCOURT
40	1186	65.1	910	6	CD794331	EST665692
41	1177	64.6	909	7	CK179258	EST768578
42	1170	64.3	928	4	BM018808	BM018808
43	1168.5	64.2	1080	6	CD461172	FG09_04p0
44	1167	64.1	896	5	BUI70979	AGENCOURT
45	1164	63.9	838	5	BUI910710	AGENCOURT

ALIGNMENTS

RESULT 1

CR617553

LOCUS

DEFINITION

full-length cDNA clone CS0D1011F14 of Placentia Cot 25-normalized

of Homo sapiens (human).

ACCESSION

CR617553

VERSION

CR617553.1 GI:50498360

KEYWORDS

HTC; CNSLT CDNA.

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 3292)

AUTHORS

Li, W.B., Gruber, C., Jessee, J., and Polayes, D.

TITLE

Full-length cDNA libraries and normalization

JOURNAL

Unpublished

REMARK

Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue

REFERENCE

2 (bases 1 to 3292)

AUTHORS

Genoscope.

TITLE

Direct Submission

JOURNAL

Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)

COMMENT

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with NotI and cloned into the NotI and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.

FEATURES

Location/Qualifiers

1..3292

/organism="Homo sapiens"

/mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D1011YF14"
 /tissue_type="placenta Cot 25-normalized"
 /plasmid="pCWSPORT_6"

ORIGIN

Alignment Scores:

Pred. No.: 2,61e-210 Length: 3292
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 3 Gaps: 0

US-10-698-228-1 (1-351) x CR617553 (1-3292)

QY 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSer 20
 Db 27 ATGGGCGACCCGGAAGCGCGAAGCGCGCTGGATCAGATCAGAGATCATCTTCA 86
 QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
 Db 87 GACACCAACGAAGTGAATAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 146
 QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
 Db 147 CGTTTGTTCATCTTCCATCCAGTACCTCTCATCTCTAGCCCTTTTTCGACGCCAGTAT 206
 QY 61 AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
 Db 207 GCTTCTCTTGACAGCAGAGAGGTGCACTTATCAAGAGATCTCCTCACTGGAACAG 266
 QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp 100
 Db 267 CTTAAAGCAGATGAGAAGTACTCTCTCATCTCTAGCCCTTTTTCGACGCCAGTAT 326
 QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 Db 327 GGAATGTAAATGAATAATTTGTGGAGCCCTTAGTCAGAGGTGAGGTTCCAGAGCT 386
 QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 Db 387 CGCTGTTCTATGCTGCTTCAATCTCATCGAGATGTTCACTCAGAGATGATACAGTTTG 446
 QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 Db 447 CTGATAGACACTTACATCAGATCCCAAGAAAGGAATTTTATTTAATGCAATTTGAA 506
 QY 161 ThrMetProTyrValLysLysLysAlaAspTyrAlaLeuArgTyrIleAlaAspArgLys 180
 Db 507 ACCATGCCCTATGTTAAGAAAGAGAGATGGGCTTTCGAGTGAATACAGATAGAAA 566
 QY 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
 Db 567 TCTACTTTTGGGAAGAGTGTGGCTTGTGCTGTAGAGAGAGTTTCTTCTCAGGA 626
 QY 201 SerPheAlaAlaIlePheThrLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 Db 627 TCTTTTGTGCTATATCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCTCTTTTCC 686
 QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 Db 687 AATGAACCTCATCAGAGATGAAGAGCTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 746
 QY 241 TyrLeuValAsnLysProSerGluArgValArgGluIleIleValAspAlaValLys 260
 Db 747 TACTTAGTAAATTAAGCTTCAGAGAAAGGGTCAGGAGATCATTTGTTGATGCTGTCAA 806
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 Db 807 ATTGACGAGGAGTTTAAACAGAGCCTTGCAGTTCGCTCATTTGGAATGAATGTCATT 866

QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 Db 867 TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTGGATTCTCA 926
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 Db 927 AAGGTTTTTTCAGGCAGAAAAATCCTTTTGTATTTATGAGAAACATTTCTTTAGAAAGGAAA 986
 QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 Db 987 ACAAAATTTCTTTGAGAAACAGATTTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC 1046
 QY 341 ThrAsnValPheThrLeuAspAlaAspPhe 351
 Db 1047 ACAGATAACGCTCTTACCTTGGATGCAGATTTT 1079

RESULT 2
 HSM802033 4748 bp mRNA linear HTC 22-SEP-2004
 LOCUS Homo sapiens mRNA; cDNA DKFZp761E1312 (from clone DKFZp761E1312).
 DEFINITION AL137348
 ACCESSION AL137348.1 GI:6807859
 VERSION HTC.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 4748)
 AUTHORS Ansorge,W., Krieger,S., Regiert,T., Rittmuller,C., Schwager,B.,
 Weses,H.W., Weill,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
 Wiemann,S.
 CONSRM The German cDNA Consortium
 TITLE Direct Submission
 JOURNAL Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764
 COMMENT Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by EMBL (European Molecular Biology Laboratories,
 Heidelberg/Germany) within the cDNA sequencing consortium of the
 German Genome Project.
 This clone (DKFZp761E1312) is available at the RZPD Deutsches
 Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
 Please contact RZPD for ordering:
 http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp761E1312
 Further information about the clone and the sequencing project is
 available at http://mips.gsf.de/projects/cdna/.

FEATURES
 Location/Qualifiers
 1..4748
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="RZPD:DKFZp761E1312"
 /db_xref="taxon:9606"
 /clone="DKFZp761E1312"
 /tissue_type="amygdala"
 /clone_lib="761 (synonym: hamy2). Vector pSport1; host
 DH10B; sites NotI + SalI"
 /dev_stage="adult"
 /note="ribonucleotide reductase"
 1..4748
 /gene="DKFZp761E1312"
 47..1102
 /gene="DKFZp761E1312"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="CAB70703.2"
 /db_xref="GI:52545612"
 /db_xref="GOA:Q9NTD8"
 /db_xref="UniProt/TREMBL:Q9NTD8"
 /translation="MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRSSRRFYI
 FFIQYPDWIKMYKQAQSFATAEVDLSKDLPHNKLKADKFI SHILAFPAASDGI
 VNEILVERFSQVQVPEARCFYGFQILINHESEMYSLIDTVIRDPKKEFLFPNATE
 TWPYVKKADWALRWIADKSTFGERVAFVAVEGVFSGSFAAIFWLKRGGLMPGLT
 FSNELISRDEGLHCDPACLMFQYLVNKPSEVRVRIIVDAVKIEQEFLEALPVGLIG

gene
 CDS

MNCILMKQVIEFVADRLLVELGFSKVFOAENPDPFMEINSLBGTNFFPKRVSEYQRF
AVMAETTDNVFTLDAF"

ORIGIN

Alignment Scores:
Pred. No.: 4,56e-210 Length: 4748
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-698-228-1 (1-351) x HSM802033 (1-4748)

QY 1 MetGluValProGluArgProGluAlaGluValLeuAspGluAspGluArgSerSer 20
DB 47 ATGGGGGACCCGGAAGGCGGAGCGCGGCTGGATCAGATCAGATCATCTTCA 106
QY 21 AspThrAsnGluSerGluLeuLeuLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
DB 107 GACACCAACGAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 166
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrPheMetTyrLysGlnAlaGln 60
DB 167 CGGTTTGTTCATCTTCCATCCAGTACCTCGATATTTGGAATGATGATAACAGGCACAG 226
QY 61 AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
DB 227 GCTTCCTCTGACAGCAGAGAGGTCGACTTATCAAGAGATCTCCCTCAGTGAACAAG 286
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaLaserAsp 100
DB 287 CTTAAGCAGATGAGAGTACTTCTCTCATCTTACCTTTTTCGAGCAGTGTAT 346
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 347 CGAATTTGAATGAAATTTGGTGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 406
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 407 CGCTGTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 466
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 467 CTGATAGACACTTATCATAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 526
QY 161 ThrMetProTyrValLysLysAlaAspTyrAlaLeuArgTyrIleAlaAspArgLys 180
DB 527 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCTTGGATGGATAGCAGATAGAAA 586
QY 181 SerThrPheGluArgValValAlaPheAlaValGluGluValPhePheSerGly 200
DB 587 TCTACTTTTGGGGAAGAGTGTGGCTTGTCTGTAGAAGGAGTTTCTTCTCAGCA 646
QY 201 SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 647 TCTTTTGTCTATATCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 706
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
DB 707 ATGAACCTCATCAGAGATGAAGGACTTCACTGTGACTTTGCTTGTCTGATGTCCAA 766
QY 241 TyrLeuValAsnLysProSerGluArgValArgGluIleValAspAlaValLys 260
DB 767 TACTTTAGTAAATAAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGTCATGTCGCAAA 826
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
DB 827 ATTGAGCAGGAGTTTAAACAGAGCCCTGGCAGTTGGCCTCATTTGGAATGAATGTCATT 886
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuGlyPheSer 300
DB 887 TTGATGAACAGTACATTTGAGTTTGTAGCTGACATTTACTTTGTGGAACCTTGATTTCTCA 946

QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
DB 947 AAGGTTTTTCAGGCAGAAATCCTTTTGATTTTATGGAAACATTTCTTTAGAGGAAA 1006
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
DB 1007 ACAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGGTTTATGCGAGAAC 1066
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB 1067 ACAGATAACGCTCTTCCCTTGGATGCGAGATTTT 1099
RESULT 3
BC042948
LOCUS
DEFINITION
Homo sapiens cDNA clone IMAGE:4798175, containing frame-shift errors.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
REMARK
COMMENT
BC042948 4650 bp mRNA linear HTC 12-OCT-2004
Homo sapiens cDNA clone IMAGE:4798175, containing frame-shift errors.
BC042948 2 GI:34194000
HTC.
Homo sapiens (human)
Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 4650)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altshul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Daplaten, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalek, U., Smalus, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 4650)
Director MGC Project.
Direct Submission
Submitted (09-JAN-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
On Aug 25, 2003 this sequence version replaced gi:27695575.
Contact: MGC help desk
Email: cgabs-x@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prabhu, Parvaneh Saeedi, JR Santos, Angeliq Schnerch, Ursula Skalska, Duane Smalus, Jeff Stott, Miranda Tsai, George Yang, Jacquie

Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E.B. Consortium/BLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 70 Row: h Column: 19
 This clone has the following problem: frame shifted.

FEATURES

source

1. .4650
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4798175"
 /tissue_type="Brain, hypothalamus"
 /clone_lib="NIH MGC_96"
 /lab_host="DH10B"
 /note="Vector: pBluescript"

ORIGIN

Alignment Scores:

Pred. No.: 2,04e-206 Length: 4650
 Score: 1791.00 Matches: 350
 Percent Similarity: 99.72% Conservative: 0
 Best Local Similarity: 99.72% Mismatches: 1
 Query Match: 98.35% Indels: 1
 DB: 3 Gaps: 0

US-10-698-228-1 (1-351) x BC042948 (1-4650)

QY 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSer 20
 DB 72 ATGGGCGACCCGGAAGCGCGAAGCGCGGCTGGATCAGATGAGATCATCTTCA 131
 QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
 DB 132 GACACCAACGAAGTGAATTAAGTCAATGAAGACCA-CTCCTAAGAAGAGTTCCTCG 190
 QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
 DB 191 CGGTTTGTCATCTTCCATCCAGTACCTGATATTTGGAAATGATATAACGCGCACAG 250
 QY 61 AlaSerPheTyrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
 DB 251 GCTTCTCTCGACAGCAGAGAGTCTGACTTATCAAGGATCTCCTCACTGGAACAAG 310
 QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp 100
 DB 311 CTAAAGCAGATGAGAAGTACTTCTCTCATCTTAGCCCTTTTTCGAGCCAGTGAT 370
 QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 DB 371 GGAATGTAAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 430
 QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 DB 431 CGCTGTTTCTATGGCTTCAATTCCTCATCGAAGTGTTCCTCATCAGAGATGTACAGTTG 490
 QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 DB 491 CTGATAGACACTTACATCAGATATCCAAAGAAAGGGAATTTTATTAATGCAATTGAA 550
 QY 161 ThrMetProTyrValLysLysValAlaAspTyrAlaLeuArgTyrIleAlaAspArgLys 180
 DB 551 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGGATGGATAGCAGATAGAAA 610
 QY 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
 DB 611 TCTACTTTTGGGGAAGAGTGGTGGCTTGTGCTGTGTAAGAGAGTTCCTCTCAGGA 670
 QY 201 SerPheAlaAlaIlePheTyrLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 DB 671 TCTTTTGTCTGCTATATCTGGCTAAAGAAAGAGAGGTCCTATGCGCAGAGCTCACTTTTCC 730
 QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240

DB 731 AATGAACATCATCAGCAGATGAAGACTTCACTGTGACTTTGCTGTGCTGATGTCCAA 790
 QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleLeuValAspAlaValLys 260
 DB 791 TACTTATTAATAAGCCTTCAGAAAGAAAGGTCAGGAGATCATTTGTGTGATGTGTCAA 850
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 DB 851 ATTGACAGAGGATTTTAAACAGAGCCTTCCAGTTGGCCTCATTTGGAATGAATTGCATT 910
 QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 DB 911 TTGATGAAACAGTACATTTAGTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 970
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 DB 971 AAGTTTTTTCAGGCAGAAAAATCCTTTTGTATTTTATGAAAAACATTTCTTTAAGAGGAAA 1030
 QY 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 DB 1031 ACAATTTCTTGGAAACAGATTTTACAGATACAGGTTTTCAGTTTATGCGAGAAACC 1090
 QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 DB 1091 ACAGATAACGCTCTTCACTTTGGATGCGAGATTTT 1123

RESULT 4
 CR627376 3501 bp mRNA linear HTC 22-SEP-2004
 LOCUS Homo sapiens mRNA; cDNA DKFZp686M05248 (from clone DKFZp686M05248) .
 DEFINITION CR627376
 ACCESSION CR627376.1 GI:50949847
 VERSION HTC.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 3501)
 AUTHORS Ottenwaelder,B., Obermaier,B., Deutschenbaur,S., Schaiipp,A.,
 Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
 Wiemann,S.
 CONSRTM The German cDNA Consortium
 TITLE Direct Submission
 JOURNAL Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764
 Neuherberg, GERMANY
 COMMENT Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by Medigenomix (Martinsried/Germany) within the cDNA
 sequencing consortium of the German Genome Project.
 This clone (DKFZp686M05248) is available at the RZPD Deutsches
 Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
 Please contact RZPD for ordering:
 http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp686M05248
 Further information about the clone and the sequencing project is
 available at <http://mips.gsf.de/projects/cdna/>.
 FEATURES
 Location/Qualifiers
 1. .3501
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKFZp686M05248"
 /tissue_type="salivary gland"
 /clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
 DH10B; sites SfiIA + SfiIB"
 /dev stage="adult"
 /note="ribonucleotide reductase M2 B (TP53 inducible)"
 1. .3501
 /gene="DKFZp686M05248"
 332. .1225
 /gene="DKFZp686M05248"
 /codon_start=1
 /product="hypothetical protein"
 gene
 CDS
 332. .1225

```

/protein_id="CAH10473.1"
/db_xref="GI:50949848"
/taxation="MYKOAQSFPTABEVDLSKTKPHNKLKADKEVFI SHILAPFAE
SDGI VNEI AVEP SQSOVPEAPCFQGIILINVHSMYSLDIITVRPKRFLP
PATETWYKKCADWALRIADRKSTFGERVVAAGA VEGFSGSFAA I PMLKRGIM
PGLTSTNMIISRDGELDCPFACMLQSTYLVNKPPEERVELIIVADAKI EQEQLFTALP
GLMGLNCLMKVOYIEFVARDLVELGFSKVPQAE NPFDMENISLEGTNPFPEKRYSE
YORFAVMAETDNVDTLADP"

```

ORIGIN

Alignment Scores:			
Pred. No.:	4.5e-200	Length:	3501
Score:	1737.50	Matches:	339
Percent Similarity:	96.06%	Conservative:	2
Best Local Similarity:	95.49%	Mismatches:	7
Query Match:	95.41%	Indels:	7
DB:	3	Gaps:	1
US-10-698-228-1 (1-351) x CR627376 (1-3501)			
QY	4	ProGluuAArgProGluuAaAlaGlyLeuAspGlnAspGlu-----	16
Db	158	CCCGCTCGCTGCTCTGCCCGGGCCAGCCAGGACGACGCTTGAGAGCCCTCAGGCGCGG	217
QY	17	ArgSerSerSerAspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuAArg	36
Db	218	AGATCATCTTCAGACACCAACGAAAGTGAATAAAGTCAANTGAGAGGCCATCTCTAAGA	277
QY	37	LysSerSerArgArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyr	56
Db	278	AAGAGTTCTCGCGGTTTGTCTATCTTCCATCCAGTACCCTCGATATTTGGAAAAATGTAT	337
QY	57	LysGlnAlaGlnAlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuPro	76
Db	338	AAACAGGCACAGGCTTCCTCTCGACACGACGAGGTCGACTTATCAAGGATCTCCCT	397
QY	77	HisTrpAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePhe	96
Db	398	CACCTGGACAAGCTTAAGCAGATGAGAGTACTTCTCATCTCTCACATCTTAGCCTTTTT	457
QY	97	AlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGln	116
Db	458	CGACGCACTGATGAATTTGTAATACAAATTTGGTGGAACGCTTTAGTCAGGAGTGCAG	517
QY	117	ValProGluAlaAArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGlu	136
Db	518	GTTCCAGAGGCTCGCTGTCTTCTATGGCTTTCAAATCTCATCAGAAATGTTCACTCAGAG	577
QY	137	MetTyrSerLeuLeuIleAspThrTyrIleAArgAspProLysLysArgGluPheLeuPhe	156
Db	578	ATGTACAGTTTGGCTGATAGACACTTACATCAGAGATCCCAGAAGAGGGAATTTTATTT	637
QY	157	AsnAlaIleGluThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIle	176
Db	638	AATGCAATTGAAACCATGCCCCATGTGTTAAGAAAAAGCAGATGGGCGCTTGGCATGGAT	697
QY	177	AlaAspArgLysSerThrPheGlyGluArgValValAlaPheAlaValGluGlyVal	196
Db	698	CGAGATGAANAATCTACTTTTTGGGGAAGAGTGGTGGCCCTTCTGCTGTAGAAGAGATT	757
QY	197	PhePheSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGly	216
Db	758	TTCTTCTCAGGATCTTTTGCTGCTATATTTCTGGCTAAAGGAAGAGAGGTCTTATGCCAGA	817
QY	217	LeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCys	236
Db	818	CTCATTTTTCCAATGAACCTCATCAGCAGAGATGAAGGACTTTCATGTGACTTTGCTTGC	877
QY	237	LeuMetPheGlnTyrLeuValAsnLysProSerGluGluArgValArgLuuIleVal	256
Db	878	CTGATGTTCCAAATCTTAGTAAATAAAGCCCTCCAGAAGAAAGGTCAGGAGATCATTTGT	937
QY	257	AspAlaValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGly	276

[illegible]

FEATURES

source

T: 9844=38T4Q/03ATT0TQ0S3=S4PI33/TT:3113:ad033S011A6:www//:q3311

Location/Qualifiers
1. .1083

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1011YF14"
/tissue_type="PLACENTA C
/clone_lib="Homo sapiens
/notes="1st strand cDNA w
primer. Five prime end
digested with Not I and
sites of the pCMVSPORT 6

```

ORIGIN

Alignment Scores:

```

Length:      1083
Matches:     319
Conservative: 0
Mismatch:    5
Indels:      4
Gaps:        0

```

5.81e-183
1591.00
97.85%
97.85%
87.37%
1

US-10-698-228-1 (1-351) x AL547501 (1-1083)

QY 1 MetGlyAspProGluArgProGluAlaAlaGly-LeuAspGlnAspGluArgSerSerSe 20

```
Db 27 ATGGCGCGCCGGAAGGCGGAGCGCGGNGCTGGATCAGGATGAGAGATCATCTTC 86
Qy 20 rAspThrAnGluSerGluIleLysSerAnGluProLeuLeuAArgLysSerSerAr 40
Db 87 AGACACCAACGAAAGTAAATGAAGTCAATGAAGAGCCACCTCTAAGAAAGAGTTCTCG 146
Qy 40 gArgPheValIlePheProIleGlnTrpProAspIleTrpLysMetTyLysGlnAlaGl 60
Db 147 CCGGTTTGTCATCTTCCTCCATCCAGTACCCTGATATTTGGAATGTAATAACAGGCACA 206
Qy 60 nAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAnLys 80
Db 207 GCCTTCCTTCGACAGCAGAGAGTGCATTCATCAAGAGTCTCCCTCAGTGAACAA 266
Qy 80 sLeuLysAlaAspGluLysTyPheIleSerHisIleLeuAlaPhePheAlaAlaSerAs 100
Db 267 GCTTAAAGCAGATGAGAAGTACTTCTATCTCTCACATCTTAGCCTTTTTCAGCCAGTGA 326
Qy 100 pGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAl 120
Db 327 TGAATTTGTAATGAATGAATTTGGTGAGCGCTTTAGTCAGGAGGTGCAGGTCCAGAGGC 386
Qy 120 aArgCysPheTyRgIyPheGlnIleLeuIleGluAsnValHisSerGluMetTyRserLe 140
Db 387 TCGCTGTTCTATGGCTTCAATCTCTCATCGAATGTTCACTCAGAGATGTACAGTTT 446
Qy 140 uLeuIleAspThrTyRlleAArgAspProLysLysArgGluPheLeuPheAsnAlaIleGl 160
Db 447 GCTGATAGACATTAATCAGATGCCAAGAAAGGGAATTTTATTAATGCAATTGA 506
Qy 160 uThrMetProTyRValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLy 180
Db 507 AACCATGCTTATGTTAAGAAAAAGCAGATGGGCTTGGCATGATGACAGATGAAGA 566
Qy 180 sSerThrPheGlyGluArgValAlaPheAlaAlaValGluGluValPhePheSerGl 200
Db 567 ATCTACTTTTGGGAAAGAGTGGTGGCTTGTCTGTAGAAGGAGTTTCTTCTCAGG 626
Qy 200 ySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSe 220
Db 627 ATCTTTTGTCTATATTTCTGGCTAAAGAAAGAGAGGTCTTATGCCAGGACTCACTTTTC 686
Qy 220 rAnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGl 240
Db 687 CAATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTGTGCTGATGTCCA 746
Qy 240 nTyRLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaVally 260
Db 747 ATACTTAGTAATTAAGCCTTCAGAAGAAAGGTTCAGGAGATCATTTGTTGATGCTGCAA 806
Qy 260 sIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIl 280
Db 807 AATTGAGCAGGAGTTTAAACAGAACCTTGCAGTGGCCTCATTTGGAATGAATTCAT 866
Qy 280 eLeuMetLysGlnTyRlleGluPheValAlaAspArgLeu-LeuValGluLeuGlyPheS 300
Db 867 TTTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTAATCTTGTGGAMCTTGGATTCT 926
Qy 300 eTyLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyL 320
Db 927 CAAGAGKTTTTCAGGAGAAATCCCTTTGATTTTATGGAAM-ATTTCCTTGAAGGAA 985
Qy 320 yeThrAsnPhePhe 324
Db 986 AAMA-AATTTCTTT 998
```

RESULT 6

AY398973

LOCUS

DEFINITION Homo sapiens HCM0069 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.

ACCESSION AY398973

```
VERSION AY398973.1 GI:39754962
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1056)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civeello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1056)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civeello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
    1..1056
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    <1..>1056
    /locus_tag="HCM0069"
gene
    1..1056
    /locus_tag="HCM0069"
ORIGIN
Alignment Scores:
Pred. No.: 3,02e-182 Length: 1056
Score: 1585.00 Matches: 313
Percent Similarity: 89.17% Conservative: 0
Best Local Similarity: 89.17% Mismatches: 38
Query Match: 87.04% Indels: 0
DB: 9 Gaps: 0
US-10-698-228-1 (1-351) x AY398973 (1-1056)
Qy 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20
Db 1 ATGGCGCGCCGGAAGGCGGAGCGCGGNGCTGGATCAGGATGAGAGATCATCTTC 60
Qy 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuAArgLysSerSerArg 40
Db 61 GACACCAACGAAAGTGAATGAAGTCAATGAAGAGCCACCTCTAAGAAAGAGTTCTCG 120
Qy 41 ArgPheValIlePheProIleGlnTrpProAspIleTrpLysMetTyRysGlnAlaGln 60
Db 121 CGGTTTGTCATCTTCCTCCATCCAGTACCCTGATATTTGGAATGTAATAACAGGCACA 180
Qy 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAnLys 80
Db 181 GCTTCTCTTCGAGCAGCAGAGAGGTGCAGTATTAACAAGGATCTCCCTCAGTGAACAA 240
Qy 81 LeuLysAlaAspGluLysTyRpheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Db 241 CTTAAAGCAGATGAGAAGTACTTCTATCTCTACATCTTAGCCTTTTTCAGCCAGTGA 300
Qy 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 301 GGAATGTAATGAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTCCAGAGGCT 360
Qy 121 ArgCysPheTyRgIyPheGlnIleLeuIleGluAsnValHisSerGluMetTyRserLeu 140
Db 361 CGCTGTTTCTATGGCTTCAATTCATCGAATGTTCACTCAGAGATGTACAGTTTG 420
Qy 141 LeuLysAspThrTyRlleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
```

```

Db 421 CTGATAGACACTTACATCAGAGATCCAGAAAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 480
Qy 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Db 481 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 540
Qy 181 SerThrPheGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
Db 541 NNNNNNNNNNGGAAAGAGTGGTGGCTTGTGCTGTAGAGAGTCTTCTCTCAGA 600
Qy 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 601 TCTTTTGTCTATATTCTGGCTTAAGAGAGAGGCTTATGTCAGCAGCACTCACTTTTCC 660
Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGGCTGATGTTCCAA 720
Qy 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAlaValLys 260
Db 721 TACTTAGTAATAAGCTTCAGAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 780
Qy 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 781 ATTGAGCAGNNNNNNNNNNNAAGCTTGGCCAGTTGGCTCATTTGGAATGAATTCATT 840
Qy 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 841 TTGATGAACACGTACATGAGTTGTGAGCTGACAGATTACTTTGTGAACTTGGATTCTCA 900
Qy 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 901 AAGGTTTTTCAGCAGAAATCTTTTGAATTTATGGAACAATTTCTTTAGAGGAA 960
Qy 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 961 ACAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTGCAGTTATGCGAAGACC 1020
Qy 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1021 ACAGATAAGCTTTCACCTTGGATGAGATTTT 1053

RESULT 7
AY398975 1043 bp DNA linear GSS 12-DEC-2003
LOCUS Mus musculus HCM0069 gene, VIRUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY398975
VERSION AY398975.1 GI:39754964
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1043)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Infering nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1043)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
TITLE Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
JOURNAL Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering

```

```

FEATURES             them based on alignment.
source               Location/Qualifiers
                    1..1043
                    /organism="Mus musculus"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:10090"
                    <1..>1043
                    /locus_tag="HCM0069"
gene
ORIGIN
Alignment Scores:
Pred. No.:          5,75e-168          Length:      1043
Score:              1468.00             Matches:      290
Percent Similarity: 86.61%              Conservative: 14
Best Local Similarity: 82.62%            Mismatches:   43
Query Match:        80.62%              Indels:       5
DB:                  9                   Gaps:         1

```

```

US-10-698-228-1 (1-351) x AY398975 (1-1043)
Qy 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
Db 1 ATGGCGCACCCGAAAGCGCCGAGCGCCGAGCGGAGAGGGGTGAGCAATTTGTTCG 60
Qy 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
Db 61 GAGACAGAAAGAAATGTGGTGAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCCGA 120
Qy 41 ArgPheValIlePhePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
Db 121 CGGTTGCTCATCTTCCATCCAGTATCTCTGATATCTGAGAAATGTAACAAGCAAGCAG 180
Qy 61 AlaSerPheTrpThrAlaGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
Db 181 GCATCTTTCTGGACAGAGAGGTTGACTTGTCAAAGGACCTCCCTCCTCTGGAACAG 240
Qy 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Db 241 CTTAAATCCGACGAGAAAGTATTTTATCTCCACATCTTAGCCTTTTTTGCAGCCAGTAT 300
Qy 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 301 GGAATTGTAATGAAACCTTGGTGGAGCGTTTGTAGTCAGGAGGTGACGTTCCAGAGCT 360
Qy 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 361 CGCTGCTTCTATGGCTTTCAGATTCTCATCAGAAATGTTCTACTCAGAGATGTACAGTTTA 420
Qy 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 421 CTAATAGACACTTACATCAGAGATCCAAAGAAAGNNNNNNNNNNNNNNNNNNNNNNNN 480
Qy 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Db 481 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 540
Qy 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
Db 541 NNNNNNNNNNGGAAAGAGTGGTGGCTTGTGCTGTGGAAGAAATTTTCTCTCTGA 600
Qy 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 601 TCATTGTGCTCAATATTCTGGCTAAAGAGAGCGCTCATGCTTGGCTGACTGACTTTTCA 660
Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGGCTGATGTTCCAA 720
Qy 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAlaValLys 260
Db 721 TACTTGTAAACAAGCCTTCAGAAGATAGAGTAGAGGAGAAATCATTTGCTGATGCTTCAA 780
Qy 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280

```



```

Db      781 ATCGAGCA-----GAAGCTTGGCTGTTGGCTCATTGGAATGAATGTGTT 827
Qy      281 LeuMetLysGlnTyrlleGluPheValAlaAspArgLeuValGluLeuGlyPheSer 300
Db      828 TTGATGAAACAGTACATGAGTTGAGCTGACAGATTACTTTGGAGAACTTGGATTCTCA 887
Qy      301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db      888 AGATTTCAGCGCAGAAAATCCCTTGATTATGAGAACATTTTCATTAGAGGAAA 947
Qy      321 ThrAsnPhePheGluLysArgValSerGluTyrlleGlnArgPheAlaValMetAlaGluThr 340
Db      948 ACAAAATTTCTTTGAGAAACGAGTTTCTGAGTATCAGCGATTGCACTCATGCCGAAACC 1007
Qy      341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db      1008 ACAGATAATGTCTTTCACCTTGGATGCAGATTTC 1040

```

RESULT 8
 CR603461 1571 bp mRNA linear HTC 21-JUL-2004
 LOCUS full-length cDNA clone CS0DH002YP23 of T cells (Jurkat cell line)
 DEFINITION of Homo sapiens (human).

ACCESSION CR603461
 VERSION CR603461.1 GI:50484268
 KEYWORDS HTC; CNSLT_CDNA.
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1571)
 AUTHORS Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600
 Faraday Avenue

REFERENCE 2 (bases 1 to 1571)
 Genoscope.
 Direct Submission
 Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
 Submitted 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)

COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen.

FEATURES
 source Location/Qualifiers
 1..1571
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DH002YP23"
 /tissue_type="T cells (Jurkat cell line)"
 /plasmid="pCMVSPORT_6"

ORIGIN

Alignment Scores:
 Pred. No.: 8,86e-167 Length: 1571
 Score: 1460.50 Matches: 277
 Percent Similarity: 88.42% Conservative: 36
 Best Local Similarity: 78.25% Mismatches: 35
 Query Match: 80.20% Indels: 6
 DB: 3 Gaps: 1

US-10-698-228-1 (1-351) x CR603461 (1-1571)

Qy 4 ProGluArgProGluAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
 Db 144 CTGAGCGGACCGCGCTCTGCCAGACGACCGCGAGAGGATCTTCAGGAGGCCAC 203

```

Qy      23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysE 38
Db      204 GGAGCCGAAAACTAAAGCAGCTGCCCGCGCGTGGAGATGAGCGCTGCTGAGAGAAA 263
Qy      38 rSerArgArgPheValIlePhePheProIleGlnTyrlleProAspIleTrpLysMetTyLysG 58
Db      264 CCCCCCGCGCTTTGTTCATCTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 323
Qy      58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLysSerLysAspLeuProHisTr 78
Db      324 GGAGAGGCTTCTCTTTGGACCGCGGAGAGGTGGACCTCTCCAAGGACATTCAGCATG 383
Qy      78 pAsnLysLeuLysAlaAspGluLysTyrlleSerHisIleLeuAlaPhePheAlaAl 98
Db      384 GGAATCCCTGAAACCCGAGGAGGAGATATTTATATCCCATGTTCTGGCTTTCTTTCAGC 443
Qy      98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
Db      444 AGCGATGGCATAGTAATAAGAACTTGGTGGAGCGATTAGCCCAAGAAATTCAGATTAC 503
Qy      118 oGluAlaArgCysPheTyrlleGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
Db      504 AGAAGCCCGCTGTTCTTATGGCTTCCAAATTTGCCATAAAGCAATACATCTGAAATGTA 563
Qy      138 rSerLeuLeuIleAspThrTyrlleArgAspProLysLysArgGluPheLeuPheAsnAl 158
Db      564 TAGTCTTCTTATTGACACTTACATAAAGATCCCAAGAAAGGAAATTTCTTCAATGC 623
Qy      158 alieGluThrMetProTyrlleValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
Db      624 CATTAAGACGATGCTTGTCTCAAGAAAGAGGAGAGCTGGGCTTGGCTGGATTGGGGA 683
Qy      178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePh 198
Db      684 CAAAGAGGCTACCTATGTTGCAACGTGTGTAGGCTTTGCTGCAGTGGAGGATTTCTT 743
Qy      198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
Db      744 TTCGGTCTTTTGGCTGATATTTCTGCTCAAGAAACGAGGAGCTGATGCTTGGCTCAC 803
Qy      218 rPheSerAsnGluLeuIleSerArgAspGluGlyLysHisCysAspPheAlaCysLeuMe 238
Db      804 ATTTCTAATGAACCTTATTAGCAGAGATGAGGGTTTACACTGTGATTGCTTGCCTGAT 863
Qy      238 tPheGlnTyrlleValAsnLysProSerGluGluArgValArgLulleIleValAspAl 258
Db      864 GTTCAAAACCTGGTACACAAACCATCGAGGAGAGAGTAAGAGAAATAATATCAATGC 923
Qy      258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
Db      924 TGTTGGATGAACAGAGGTTCTCTACTGAGGCTTGCCTGTGAGGCTCATTTGGGATGA 983
Qy      278 nCysIleLeuMetLysGlnTyrlleGluPheValAlaAspArgLeuLeuValGluLeuGl 298
Db      984 TTGCACCTTAATGAAGCAATACATTGATTGTTGGCAGACAGACTTATGCTGGAACCTGG 1043
Qy      298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGl 318
Db      1044 TTTTAGCAGGTTTTCAGAGTAGAGAACCCCATTTTGACTTTTATGGAGATATTTTCACTGA 1103
Qy      318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrlleGlnArgPheAlaValMetAl 338
Db      1104 AGGAAAGACTTAATCTTTTGAAGAGAGAGTAGGCGGAGTATCAGAGGATGGGAGTGTGTC 1163
Qy      338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db      1164 AAGTCCACAGAGAAATTTCTTTTACCTTGGATGCTGACTTC 1203

```

RESULT 9
 CR625489
 LOCUS full-length cDNA clone CS0DA003Y109 of Neuroblastoma of Homo
 DEFINITION sapiens (human).

1573 bp mRNA linear HTC 21-JUL-2004
 full-length cDNA clone CS0DA003Y109 of Neuroblastoma of Homo
 sapiens (human).

ACCESSION CR625489
 VERSION CR625489.1 GI:50506296
 KEYWORDS HTC; CNSLT cDNA.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1573)
 AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue
 REFERENCE 2 (bases 1 to 1573)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
 COMMENT 1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
 FEATURES
 source Location/Qualifiers
 1..1573
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSODA003YL09"
 /tissue_type="Neuroblastoma"
 /plasmid="pCMVSPORT_6"
 ORIGIN
 Alignment Scores:
 Pred. No.: 8, 87e-167 Length: 1573
 Score: 1460.50 Matches: 277
 Percent Similarity: 88.42% Conservative: 36
 Best Local Similarity: 78.25% Mismatches: 35
 Query Match: 80.20% Indels: 6
 DB: 3 Gaps: 1
 US-10-698-228-1 (1-351) x CR625489 (1-1573)
 QY 4 ProGluArgProGluAlaAlaGlyLeuAspGluArgSerSerSerAspThrAs 23
 DB 155 CTGAGCGGACCGCGCTCTGGCCAGCAAGACCGGAGGAGGATCTCCAGGAGCCAC 214
 QY 23 nGluSerGluLeuLysSer-----AsnGluGluProLeuLeuArgLysSe 38
 DB 215 GGAGCCGAAACTAAAGCAGCTGCCCGCGGTGGAGGATGAGCGCTCTGAGAGAAA 274
 QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysG 58
 DB 275 CCCCCGCGCTTTCATCTCCCATCGATGACCATGATATCTGCAGATGATAGAA 334
 QY 58 nAlaGlnAlaSerPheThrThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
 DB 335 GGCAGAGGCTCTCTTTGGACCGCGGAGGAGTGACCTCTCCAGGACATTCAGCAGT 394
 QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAl 98
 DB 395 GGAATCCCTGAAACCGGAGAGAGATATTTATATCCCATCTGCTGCTTCTTTCGAGC 454
 QY 98 aserAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
 DB 455 AAGCATGGCATAGTAATGAAACTTGGTGAGGCAATTTAGCCAGAGATTCAGATTAC 514
 QY 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
 DB 515 AGAAGCCCGCTGTTCTATGCTTCCAAATGCCATGGAAACATACATCTTGAATGTA 574

QY 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
 DB 575 TAGCTTCTTATTGACACTTACATAAAGATCCCAAGAAAGGAATTTCTTCAATGC 634
 QY 158 alIeGluThrMetProTyrValLysLysLysAlaAspTyrAlaLeuArgTyrIleAlaAs 178
 DB 635 CATTAAGACGATCGCTTGTCTCAAGAAAGAGGAGACTGGCCCTTGCCTGGATTGGGA 694
 QY 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPheP 198
 DB 695 CAAAGAGGCTACCTATGTTGAACGTGTGTAGCTTTTGTCTCAGTGGAGGCAATTTCTT 754
 QY 198 eSerGlySerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuTh 218
 DB 755 TTCGGTCTTTTTCGTCGATATTCGCTCAAGAAACGAGACTGATGCTGGCCTCAC 814
 QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
 DB 815 ATTTCTAATGAACCTTATTAGCAGAGATGAGGGTTTACACTGTGATTTGCTTGCCTGAT 874
 QY 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
 DB 875 GTTCAACACCTCGTACAAACCATCGGAGGAGAGATAGAGAAATAATATCAATGC 934
 QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
 DB 935 TGTTCCGATAGAACAGAGTTCTCTACTGAGGCTTTCGCTGTGAGCTCATTTGGATGAA 994
 QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGl 298
 DB 995 TTGCACCTCTAATGAAGCAATACATTTAGTTTGTGGCAGACAGACTTATGTCTGGAAC 1054
 QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGl 318
 DB 1055 TTTTAGCAAGGTTTTCAGAGTAGAGAACCCATTTGACTTTATGGAGATATTTTCACTGA 1114
 QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
 DB 1115 AGGAAGACCTAATCTTTGAGAGAGAGATGAGCGAGTATCAGAGATGGGAGTGTATGTC 1174
 QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
 DB 1175 AAGTCCACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1214
 RESULT 10
 LOCUS CR608076
 DEFINITION full-length cDNA clone CS0DK012YD23 of HeLa cells Cot 25-normalized of Homo sapiens (human).
 ACCESSION CR608076
 VERSION CR608076.1 GI:50488883
 KEYWORDS HTC; CNSLT cDNA.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1582)
 AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue
 REFERENCE 2 (bases 1 to 1582)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
 COMMENT 1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a

```

division of Invitrogen.
FEATURES             Location/Qualifiers
     source          1..1582
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DA011YH23"
                     /tissue_type="HeLa cells Cot 25-normalized"
                     /plasmid="pCMVSPORT_6"

ORIGIN
Alignment Scores:
Pred. No.:          8.95e-167      Length:          1582
Score:              1460.50        Matches:          277
Percent Similarity: 88.42%         Conservative:    36
Best Local Similarity: 78.25%      Mismatches:     35
Query Match:        80.20%         Indels:          6
DB:                  3              Gaps:             1

US-10-698-228-1 (1-351) x CR608076 (1-1582)
QY      4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
Db      142 CCTGAGCGGACCCGCGTCTGCGCAGCAAGACCGCGAGGAGGATCTTCCAGGAGCCAC 201
QY      23 nGluSerGluIleLysSer-----AenGluGluProLeuLeuArgLysSe 38
Db      202 GGAGCGGAAGTAAGCAGCTGCCCGCGGTGGAGGATGAGCCGCTCTGAGAGAAA 261
QY      38 rSerArgArgPheValIlePheProIleGlnTyPrAspIleTyrLysMetTyrLysG 58
Db      262 CCCCCGCGCTTGTCACTCTCCCATCGATGATGATGATGATGATGATGATGATGATGAT 321
QY      58 nAlaGlnAlaSerPheTyrThrAlaGluGluValAlaAspLeuSerLysAspLeuProHisTr 78
Db      322 GGCAGAGGCTCTTTGGACCGCGCGAGGAGTGACCTCTCAAGAGCATTGAGCAGCTG 381
QY      78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAla 98
Db      382 GGAATCCCTGAAACCGGAGGAGATATTTATATCCCATGTTCTGGCTTTCTTTCAGC 441
QY      98 sSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
Db      442 AAGCATGGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 501
QY      118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleLeuValAlaHisSerGluMetTy 138
Db      502 AGAAGCCGCTGTTCTATGGCTTCCAAATTCATGCGAAGGAGGAGGAGGAGGAGGAGGAG 561
QY      138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
Db      562 TAGTCTCTTATGTGACCTTACATATAAGATCCCAAGAAAGGAAATTTCTCTTCAATGC 621
QY      158 aIleGluThrMetProTyValLysLysLysAlaAspTyrAlaLeuArgTyrIleAlaAs 178
Db      622 CATGGAACGATGCTTGTGTCAAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 681
QY      178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaAlaValAlaGluGlyValPhePh 198
Db      682 CAAAGAGGCTACTATGTTGTAAGCTGTGTAGCTTTTGTGCTGCTGGAAGGCAATTTCTT 741
QY      198 eSerGlySerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuTh 218
Db      742 TTCGCGTCTTTTGGCTGATATTTCTGCTCAAGAAACGAGGAGGAGGAGGAGGAGGAGGAG 801
QY      218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
Db      802 ATTTCTAATGACCTTATTAGCAGAGATGAGGTTTACACTGTGATTTTGTCTTGCCTGAT 861
QY      238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
Db      862 GTTCAACACCTGGTACACAAACCATCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 921

```

```

QY      258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
Db      922 TGTTCGATAGAACAGAGGTTCTCTACTGAGGCGCTTGCCTGTGAAGCTCATTTGGATGAA 981
QY      278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGl 298
Db      982 TTGCACCTCTAATAGCAATACATTTGAGTTTGTGGCAGACAGACTTATGCTTGGAACTGGG 1041
QY      298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGl 318
Db      1042 TTTTAGCAAGGTTTTCAGATGATAGAACCCCATTTGACTTTATGGAGATAATTTTCACTGGA 1101
QY      318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
Db      1102 AGGAAGACTAATCTTTTGGAGAGAGAGTAGGCGAGTATCAGAGGATGGGAGTGTATGTC 1161
QY      338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db      1162 AAGTCCAACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1201

RESULT 11
CR602054          1588 bp      mRNA      linear      HTC 21-JUL-2004
LOCUS             full-length cDNA clone CS0DA011YH14 of Neuroblastoma of Homo
DEFINITION        sapiens (human).
ACCESSION         CR602054
VERSION           CR602054.1 GI:50482861
KEYWORDS          HTC; CNSLIT_CDNA.
SOURCE            Homo sapiens
ORGANISM          Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE          1 (bases 1 to 1588)
AUTHORS           Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE             Full-length cDNA libraries and normalization
JOURNAL           Unpublished
REMARK            Contact : Peng Liang Email : fliang@lifetech.com URL :
                  http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
                  Paraday Avenue
                  2 (bases 1 to 1588)
REFERENCE          Genoscope.
AUTHORS           Direct Submission
TITLE             Submitted (20-JUL-2004) Genoscope - Centre National de Sequençage :
JOURNAL           BP 131 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr)
COMMENT            - Web : www.genoscope.cns.fr)
                  1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
                  end enriched, double-strand cDNA was digested with Not I and cloned
                  into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
                  was normalized. Library was constructed by Life Technologies, a
                  division of Invitrogen.
FEATURES           Location/Qualifiers
     source        1..1588
                   /organism="Homo sapiens"
                   /mol_type="mRNA"
                   /db_xref="taxon:9606"
                   /clone="CS0DA011YH14"
                   /tissue_type="Neuroblastoma"
                   /plasmid="pCMVSPORT_6"

ORIGIN
Alignment Scores:
Pred. No.:          9e-167      Length:          1588
Score:              1460.50        Matches:          277
Percent Similarity: 88.42%         Conservative:    36
Best Local Similarity: 78.25%      Mismatches:     35
Query Match:        80.20%         Indels:          6
DB:                  3              Gaps:             1

US-10-698-228-1 (1-351) x CR602054 (1-1588)
QY      4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
Db      142 CCTGAGCGGACCCGCGTCTGCGCAGCAAGACCGCGAGGAGGATCTTCCAGGAGCCAC 201

```

```
QY 23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
Db 202 GGAGCGGAAACTAAAGCAGCTGCCCGCGCTGGAGGATGAGCGCTGCTGAGAGAAA 261
QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysG1 58
Db 262 CCCCCCGCGCTTTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGATATAAGAA 321
QY 58 nAlaGlnAlaSerPheTyrThrAlaGluLeuValAspLeuSerLysAspLeuProHisTr 78
Db 322 GGACAGAGGCTTCTTTTGGACCGCCGAGGAGTGGACCTCTCAAGACACATTCAGCACGTG 381
QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAl 98
Db 382 GGAATCCCTGAAACCGAGGAGGATATTTATATCCATGTTCTGGCTTCTTTCGAGC 441
QY 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
Db 442 AAGCGATGCATAGTAATGAAACCTTGGTGGAGGATTTAGCAAGAGTTCAGATTAC 501
QY 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
Db 502 AGAAGCCCGCTGTTCTATGGCTTCCAAATTGCCATGGAAACATACATTCGAAATGTA 561
QY 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
Db 562 TAGTCTTCTTATGACACTTACATTAAGATCCCAAGAAAGGGAATTTCTCTCAATGC 621
QY 158 alLeGluThrMetProTyrValLysLysLysAlaAspTyrAlaLeuArgTyrIleAlaAs 178
Db 622 CATTGAAACGATGCTTGTGTCAAGAAAGAGCAGACTGGCCCTTGGCTGGATGGGA 681
QY 178 pArgLysSerThrPheGlyGluArgValIleAlaPheAlaValGluGlyValPhePh 198
Db 682 CAAAGAGGCTACTATGTGTGAACGTGTGTAGCTTGTGCTGAGTGGAGGCAATTTCTT 741
QY 198 eSerGlySerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuTh 218
Db 742 TTCGGTCTTCTTTGGCTGATATCTTGGCTCAAGAAACGAGGACTGATGCTGGCTCAC 801
QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
Db 802 ATTTCTAATGAATATTATGACAGATGAGGGTTTACACTGTGATTTGCTTGCCTGAT 861
QY 238 tPheGlnTyrLeuValAsnLysProSerGluArgValArgGluIleIleValAspAl 258
Db 862 GTTCAAAACACCTGTTACACAAACCATCGAGGAGAGTAAAGAAATAATATCAATGC 921
QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
Db 922 TGTTCGGATAGAACAGGAGTTCTCACTAGAGCTTGGCTGTGAAGCTCATTTGGGATGAA 981
QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuG1 298
Db 982 TTGCACCTCTAATGAGCAATACATTTGAGTTTGTGGCAGACAGACTTATGCTGGAATGGG 1041
QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG1 318
Db 1042 TTTTAGCAAGGTTTTCAGAGTAGAGAACCCATTTGACTTTATGGAGATATTTCACTGGA 1101
QY 318 uGlyLysThrAnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetal 338
Db 1102 AGGAAGACTACTCTTTTGAAGAGAGTAGGCGAGTATCAGAGATGGGATGTGATGTC 1161
QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1162 AAGTCAACAGAGAAATTTCTTTTACCTTGTGATGCTGACTTC 1201
RESULT 12
CR621427 1592 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CS0DE011YL04 of Placenta of Homo sapiens
DEFINITION
```

```

(human).
CR621427
VERSION CR621427.1 GI:50502234
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1592)
AUTHORS Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
Faraday Avenue
REFERENCE 2 (bases 1 to 1592)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source
location/Qualifiers
1..1592
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE011YL04"
/tissue_type="Placenta"
/plasmid="pCMVSPORT_6"

Alignment Scores:
Pred. No.: 9,04e-167 Length: 1592
Score: 1460.50 Matches: 277
Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
DB: 3 Gaps: 1

US-10-698-228-1 (1-351) x CR621427 (1-1592)
QY 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
Db 153 CTTGAGCGGAGCCCGCTCTGCGCAGCAGACCGCGAGGAGGATCTTCCAGGAGCCAC 212
QY 23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
Db 213 GGAGCGGAAACTAAAGCAGCTGCCCGCGCTGGAGGATGAGCGCTGCTGAGAGAAA 272
QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysG1 58
Db 273 CCCCCCGCGCTTTGTCATCTTCCCATCGATCATCATGATATCTGGCAGATGATATAAGAA 332
QY 58 nAlaGlnAlaSerPheTyrThrAlaGluLeuValAspLeuSerLysAspLeuProHisTr 78
Db 333 GGACAGAGGCTTCTTTTGGACCGCCGAGGAGTGGACCTCTCAAGACATTCAGCACGTG 392
QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAl 98
Db 393 GGAATCCCTGAAACCGAGGAGGATATTTATATCCATGTTCTGGCTTCTTTCGAGC 452
QY 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
Db 453 AAGCGATGCATAGTAATGAAACCTTGGTGGAGGATTTAGCAAGAGTTCAGATTAC 512
QY 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
Db 513 AGAAGCCCGCTGTTCTATGGCTTCCAAATTGCCATGGAAACATACATTCGAAATGTA 572
```

```

Qy 138 rSerLeuLeuIleAspThrTyrIleArcAspProLysLysArgGluPheLeuPheAsnAl 158
Db 573 TAGTCTCTTTTACACTTACATGAAGATCCCAAGAGGAAGGAAATTTCTTTCAATGC 632
Qy 158 aileGluThrMetProTyrValLysLysAlaAspTrrAlaLeuArgTrrPileAlaAs 178
Db 633 CATTGAACGATGCTTGTGTCAAGAAGGACAGCTGGGCTTGGCTGGATGGGA 692
Qy 178 pArgLysSerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePh 198
Db 693 CAAAGAGGCTACCTATGCTGAACGTTGTGTACCTTTGTCTGAGTGAAGGCAATTTCTT 752
Qy 198 eSerGlySerPheAlaAlaIlePheTrrPheLysLysArgGlyLeuMetProGlyLeuTh 218
Db 753 TTCGGTCTTTTGGCTCGATATTCCTCAAGAAACGAGGACTGATGCTGGCTCAC 812
Qy 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
Db 813 ATTTCTAATGAACCTATTAGCAGATGAGGGTTTACACTGTGATTTTGTCTGCTGAT 872
Qy 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluLeuIleValAspAl 258
Db 873 GTTCAACACCTGTGTACACAAACCATCGAGGAGAGAGTAAGAGAAATAATTATCAATGC 932
Qy 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
Db 933 TGTTCGGATAGAACAAGGAGTTCCTCACTGAGGCTTGGCTGTGAAGCTCATTTGGGATGA 992
Qy 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuG 298
Db 993 TTGCACTCTAATGAACATAATGAGTTTGTGGCAGACACTATGCTGGAAGTGG 1052
Qy 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG 318
Db 1053 TTTTAGCAAGGTTTTCAGAGTAGAGAACCCATTTCACATTATGAGAAATATTTCACCTGA 1112
Qy 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
Db 1113 AGGAAGACTAATCTTTTGAAGAGAGTAGGCGAGTATCAGAGATGAGGAGTGTATGTC 1172
Qy 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1173 AAGTCCACACAGAAATCTTTTACCTTGGATGCTGACTTC 1212

RESULT 13
CR604378 1600 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CS0DI037Y003 of Placenta Cot 25-normalized
DEFINITION of Homo sapiens (human).
ACCESSION CR604378
VERSION CR604378.1 GI:50485185
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1600)
Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
Genoscope.
2 (bases 1 to 1600)
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library

```

was normalized. Library was constructed by Life Technologies, a division of Invitrogen.

FEATURES

source Location/Qualifiers
 1..1600
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DI037Y003"
 /tissue_type="Placenta Cot 25-normalized"
 /plasmid="pCMVSPORT_6"

ORIGIN

Alignment Scores:

Pred. No.: 9,11e-167 Length: 1600
 Score: 1460.50 Matches: 277
 Percent Similarity: 88.42% Conservative: 36
 Best Local Similarity: 78.25% Mismatches: 35
 Query Match: 80.20% Indels: 6
 DB: 3 Gaps: 1

US-10-698-228-1 (1-351) x CR604378 (1-1600)

```

Qy 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
Db 155 CCTGAGCGGACCCGCGTCTCGCCAGCAAGACCGCGGAGGAGTCTTCCAGGAGCCAC 214
Qy 23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
Db 215 GGAGCCGAAACTTAAAGACAGCTCCCGCGCGTGGAGGATGAGCGCTGTGAGAGAAA 274
Qy 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTrrPlysMetTyrLysG 58
Db 275 CCCCCCGCGTTTGTCTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGA 334
Qy 58 nAlaGlnAlaSerPheTrrPthrAlaGluGluValAlaAspLeuSerLysAspLeuProHis 78
Db 335 GGCAGAGGCTTCTTTTGGACCGCCGAGGAGGTGGACCTCTCCAAGGACATTCAGCACTG 394
Qy 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAl 98
Db 395 GGAATCCCTGAAACCCGAGGAGAGATATTTATATCCATGTCTTCTGGCTTTCTTTG 454
Qy 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
Db 455 AACCGATGGCATAGTAAATGAAACCTTGGTGGAGCGGATTTAGCCAGAGATTCAGATTAC 514
Qy 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
Db 515 AGAAGCCCGCTGTTCTATGCTTCCAAATTTGCCATGGAACACATACATTTCTGAAATGTA 574
Qy 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
Db 575 TAGTCTTCTTATTGACACTTACATAAAGATCCCAAGAGAGGAATTTCTTCTCAATGC 634
Qy 158 alieGluThrMetProTyrValLysLysAlaAspTrrAlaLeuArgTrrPileAlaAs 178
Db 635 CATTGAACGATGCTTGTGTCAAGAAAGAGGACAGACTGGGCTTGGCTGATTTGGGA 694
Qy 178 pArgLysSerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePh 198
Db 695 CAAAGAGGCTACCTATGCTGAACGTTGTGTAGCCTTTTGTCTGCTGAGGAGCAATTTCT 754
Qy 198 eSerGlySerPheAlaAlaIlePheTrrPheLysLysArgGlyLeuMetProGlyLeuTh 218
Db 755 TTCGGTCTTTTGGCTCGATATTTGGCTCAAGAAACGAGGACTGATGCTGCTGCCTCAC 814
Qy 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
Db 815 ATTTCTAATGAACCTATTATAGCAGAGATGAGGGTTTACACTGTGATTTTGTGCTGCTGAT 874
Qy 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
Db 875 GTTCAACACCTGTGTACACAAACCATCGGAGGAGAGTAGTAAAGAGAAATAATTATCAATGC 934

```

```

QY 258 aVallyleLeuGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
Db 935 TGTTGGATAGAACAGAGTTCCTACTGAGGCTTGCTGTGTAAGCTCATTTGGATGAA 994
QY 278 nCysIleLeuMetLysGlnTyrlleGluPheValAlaAspArgLeuLeuValGluLeuG1 298
Db 995 TTGCACCTTAATGAAGCAATACATTGAGTTTGTGGCAGACAGACTTATGCTGGAACCTGG 1054
QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG1 318
Db 1055 TTTTACAGGTTTTCAGAGTAGAACCAATTTGACTTTATGGAGAAATATTTTCACTGGA 1114
QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrlGlnArgPheAlaValMetAl 338
Db 1115 AGGAAGACTACTTCTTTGAGAGAGAGTAGGCGAGTAGATCAGAGATGGAGTAGTGC 1174
QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1175 AAGTCCACACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1214

RESULT 14
CR590959
LOCUS full-length cDNA clone CS0DE013YD22 of Placenta of Homo sapiens
DEFINITION (human).
ACCESSION CR590959
VERSION CR590959.1 GI:50471766
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1605)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
Genoscope.
REFERENCE 2 (bases 1 to 1605)
AUTHORS Direct Submission
TITLE Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
source
1..1605
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE013YD22"
/tissue_type="Placenta"
/plasmid="pCMVSPORT_6"

ORIGIN
Alignment Scores:
Pred. No.: 9,15e-167 Length: 1605
Score: 1460.50 Matches: 277
Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
DB: 3 Gaps: 1

US-10-698-228-1 (1-351) x CR590959 (1-1605)

QY 4 ProGluArgProGluAlaGlyLeuAspGlnAspGlu-ArgSerSerAspThrAs 23
Db 1175 AAGTCCACACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1209

RESULT 15
CR614990
LOCUS CR614990

```

```

Db 150 CTTGAGCGGACCCGCGTCTTGGCCAGCACGCCGAGGAGGATCTTTCAGGAGCCAC 209
QY 23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
Db 210 GGAGCCGAAACCTAAAGCAGCTGCCCGCGGTGGAGGATGAGCGCTGCTGAGAGAAA 269
QY 38 rSerArgArgPheValIlePhePheProIleGlnTyrlProAspIleTrpLysMetTyrlGlyG1 58
Db 270 CCCC CGCGCTTTGTGTCATCTTCCCATCGAGTACCATGATATCTGCAGATGCTATAAGAA 329
QY 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
Db 330 GGCAGCGCTTCTTTTGGACCGCGAGGAGGTGGACCTCTCCAAGGACATTCAGCACTG 389
QY 78 pAsnLysLeuLysAlaAspGluLysTyrlPheIleSerHisIleLeuAlaPhePheAlaAl 98
Db 390 GGAATCCCTGAACCCCGAGGAGAGATATTTATATCCCATGCTTCTGGCTTTCTTTCGACG 449
QY 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
Db 450 AAGCGATGGCATAGTAATGNAACCTTGGTGGAGCGATTTAGCCAGAGATTCAGATTAC 509
QY 118 oGluAlaArgCysPheTyrlGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
Db 510 AGAAGCGCGCTGTTCTATGGCTTCCAAATTTGCCATGGAAAAACATACATCTCTGAAATGTA 569
QY 138 rSerLeuLeuLeuAspThrTyrlleArgAspProLysLysArgGluPheLeuPheAsnAl 158
Db 570 TAGTCTTCTTATTGACACTTACATAAAGATCCCAAGAGAAAGGGAATTTCTTCCAATGC 629
QY 158 aIleGluThrMetProTyrlValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
Db 630 CATTGAACCATGCTTGTGTCAAGAAAGAGGAGAGCTGGGCTTTCCTGATTTGGGA 689
QY 178 pArgLysSerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePh 198
Db 690 CAAAGAGGCTACCTATGTTGTAACGCTTGTAGCTTTCCTCAGTGGAGGCAATTTCTT 749
QY 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
Db 750 TTCCGGTTCTTTTGGCTCGATATTTCTGGCTCAAGAAACGAGGACTGTGCTGGCTCAC 809
QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
Db 810 ATTTTCTAATGAACCTTATAGCAGATAGAGGTTTACACTGTGATTTGCTGCTGAT 869
QY 238 tPheGlnTyrlLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
Db 870 GTTCAACACCTGGTACACAAACCATCGAGGAGAGAGTAGAGAAATAATATCAATGC 929
QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
Db 930 TGTTCCGATAGAACAGAGGTTCTCTACTGAGGCTTGCCTGTGAAGCTCATTTGGATGAA 989
QY 278 nCysIleLeuMetLysGlnTyrlleGluPheValAlaAspArgLeuLeuValGluLeuG1 298
Db 990 TTGCACCTTAATGAAGCAATACATTGAGTTTGTGGCAGACAGACTTATGCTGGAACCTGG 1049
QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG1 318
Db 1050 TTTTACAGGTTTTCAGAGTAGAACCAATTTGACTTTATGGAGAAATATTTTCACTGGA 1109
QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrlGlnArgPheAlaValMetAl 338
Db 1110 AGGAAGACTACTTCTTTCAGAAAGAGAGTAGGCGAGTATCAGAGGATGGAGTAGTGC 1169
QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1170 AAGTCCACACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1209

RESULT 15
CR614990
LOCUS CR614990

```

CR614990 1612 bp mRNA linear HTC 21-JUL-2004

```

DEFINITION full-length cDNA clone CS0DA009YB23 of Neuroblastoma of Homo
             sapiens (human).
ACCESSION   CR614990
VERSION     CR614990.1 GI:50495797
KEYWORDS    HTC; CNSLT_cDNA.
SOURCE      Homo sapiens
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 1612)
AUTHORS     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE       Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL     1 (bases 1 to 1612)
REMARK      Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
            Full-length cDNA libraries and normalization
            Unpublished
            Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
            Faraday Avenue
REFERENCE   2 (bases 1 to 1612)
AUTHORS     Genoscope.
TITLE       Direct Submission
JOURNAL     Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
FEATURES             Location/Qualifiers
     source           1..1612
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DA009YB23"
                     /tissue_type="Neuroblastoma"
                     /plasmid="pCMVSPORT_6"
ORIGIN
Alignment Scores:
Pred. No.:          9,21e-167      Length:          1612
Score:              1460.50         Matches:         277
Percent Similarity: 88.42%          Conservative:    36
Best Local Similarity: 78.25%       Mismatches:       35
Query Match:        80.20%          Indels:          6
DB:                  3              Gaps:             1

US-10-698-228-1 (1-351) x CR614990 (1-1612)
QY      4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
Db      152 CCTGAGCGGACCGCGCTCTGGCCAGCAAGACCGCGAGGAGTCTTCCAGGAGCCAC 211
QY      23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
Db      212 GGAGCCGAAACTAAAGCAGCTGCCCGCGCTGGAGGATGAGCCGCTGCTGAGAGAAA 271
QY      38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTrrPlyMetTyrLysGl 58
Db      272 CCCCCCGCGCTTGTTCATCTCCCATCCAGTACCATGATATCTGCAGATGTATAAGAA 331
QY      58 nAlaGlnAlaSerPheTrrPThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
Db      332 GCAGAGGCTTCTTTTGGACCGCGAGGAGGTGGACCTCTCCAGAGACATTCCAGCCTG 391
QY      78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAl 98
Db      392 GGAATCCCTGAAACCGAGAGAGATATTTATATCCATGTTCTGCTTCTTTCAGC 451
QY      98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
Db      452 AAGCGATGCATAGTAAATGAAACCTTGGTGGAGCGATTAGCCAAAGAGTTCCAGATTAC 511
QY      118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138

```

```

Db      512 AGAAGCCCGCTGTTTCTATGGCTTCCAAATGGCCATGGAACATACATCTCTGAAATGTA 571
QY      138 rSerLeuLeuLeuAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
Db      572 TAGTCTCTTTATTGACACTTACATAAAGATCCCAAAGAAAGGGAATTTCTTCCAATGC 631
QY      158 aIleGluThrMetProTyrValLysLysLysAlaAspTrrAlaLeuArgTrrPilleAlaAs 178
Db      632 CATTGAAACGATCGCTTGTTGTCGAAGAAGAGCAGACTGGGCCCTTGGCTGGATTGGGA 691
QY      178 pArgLysSerThrPheGlyGluArgValAlaPheAlaAlaValGluGluValPhePh 198
Db      692 CAAAGAGGCTACCTGCTGAACGCTGTGTAGCTTTGCTGCTGAGGAGCATTTTCTT 751
QY      198 eSerGlySerPheAlaAlaIlePheTrrPheLysLysArgGlyLeuMetProGlyLeuTh 218
Db      752 TTCGGGTTCTTTTGGCTCGATATTTCTGGCTCAAGAAACGAGGACTGATGCTTGGCCTCAC 811
QY      218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
Db      812 ATTTCTAATGAACCTTATTAGCAGAGATGAGGGTTTACACTGTGATTTTGTGCTGCTGAT 871
QY      238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
Db      872 GTTCAAAACACCTGCTGACACAAACCATCGAGGAGAGATAGAGAAAATAATTATCAATGC 931
QY      258 aValLysIleGluGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
Db      932 TGTTGGATAGAACAGAGATTCTCTACTGAGGCCCTTGCTGTGAAGCTCATTTGGGATGAA 991
QY      278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluGlu 298
Db      992 TTGCACCTCTAATCGACGATACATTGAGTTTGTGGCAGACAGACTTATGCTGGAACCTGG 1051
QY      298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGl 318
Db      1052 TTTTAGCAAGGTTTTCAGAGTAGAGAACCCATTGACTTTTATGAGAAATATTTCACTGGA 1111
QY      318 uGlyLysThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
Db      1112 AGGAAAGACTAACTTCTTTGAGAAAGAGAGATGAGCGAGATATCAGAGGATGGGAGTGATGTC 1171
QY      338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db      1172 AAGTCCACACAGAGAAATTTCTTTTACCTTGGATGCTGACTTC 1211

```

Search completed: October 30, 2005, 08:53:08

Job time : 2604 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 30, 2005, 07:02:47 ; Search time 156 Seconds
(without alignments)

3681.623 Million cell updates/sec

Title: US-10-698-228-1

Perfect score: 1821

Sequence: 1 MGDPERPEAAGLDQDERSS.....QRFVMAETTDNVFTLDADF 351

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO spool/US10698228/runat_26102005_100715_5137/app_query.fasta_1_519
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DLOCALIGN=200 -THR SCORE=ptc -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10698228 @cgn_1_54 @runat_26102005_100715_5137 -NCPUS=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:**

1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/prodata/1/ina/PCUTUS_COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1821	100.0	1053	US-10-019-733-2	Sequence 2, Appli
2	1821	100.0	1053	US-10-019-733-12	Sequence 12, Appli
3	1821	100.0	1081	US-10-019-733-4	Sequence 4, Appli
4	1821	100.0	4955	US-10-019-733-3	Sequence 3, Appli
5	1460.5	80.2	2479	US-09-949-016-2025	Sequence 2025, Ap
6	1460.5	80.2	2500	US-09-962-665-9	Sequence 9, Appli
7	1460.5	80.2	2500	US-09-023-655-1370	Sequence 1370, Ap
8	1460.5	80.2	2500	US-09-963-333-9	Sequence 9, Appli
9	1460.5	80.2	2500	US-09-962-677-9	Sequence 9, Appli
10	1460.5	80.2	2500	US-09-949-016-145	Sequence 145, Appli
11	1210	66.4	14176	US-08-307-499-1	Sequence 1, Appli
12	1210	66.4	14176	US-08-307-499-14	Sequence 14, Appli

C	13	1210	66.4	14176	3	US-09-299-268-1	Sequence 1, Appli
	14	1210	66.4	14176	3	US-09-299-268-14	Sequence 14, Appli
	15	1131	62.1	1236	4	US-09-248-796A-3914	Sequence 3914, Ap
	16	1072.5	58.9	825	1	US-08-307-499-40	Sequence 40, Appli
	17	1072.5	58.9	825	3	US-09-299-268-40	Sequence 40, Appli
	18	984.5	54.1	1112	1	US-08-136-743B-1	Sequence 1, Appli
	19	969.5	53.2	1083	4	US-09-248-796A-3915	Sequence 3915, Ap
	20	768	42.2	481	3	US-08-905-223-125	Sequence 125, Ap
	21	674	37.0	656	3	US-09-328-111-791	Sequence 791, Ap
	22	599	32.9	449	4	US-09-270-767-14670	Sequence 14670, A
	23	558	30.6	419	3	US-08-905-223-129	Sequence 129, App
	24	543.5	29.8	11820	4	US-09-949-016-11887	Sequence 11887, A
	25	543.5	29.8	11826	4	US-09-949-016-13767	Sequence 13767, A
	26	513.5	28.2	13706	4	US-09-902-540-1124	Sequence 1124, Ap
	27	509.5	28.0	1053	4	US-09-902-540-2666	Sequence 2666, Ap
	28	488.5	26.8	366	1	US-08-307-499-13	Sequence 13, Appli
	29	488.5	26.8	366	3	US-09-299-268-13	Sequence 13, Appli
	30	392.5	21.6	297	4	US-09-313-294A-4435	Sequence 4435, Ap
	31	381.5	21.0	1338	4	US-09-328-352-588	Sequence 588, Ap
	32	378.5	20.8	1707	4	US-09-252-991A-3180	Sequence 3180, Ap
	33	377.5	20.7	1431	4	US-09-252-991A-2787	Sequence 2787, Ap
	34	370	20.3	984	4	US-09-252-991A-2692	Sequence 2692, Ap
	35	355	19.5	373	4	US-09-401-064-319	Sequence 319, App
	36	355	19.5	373	4	US-09-401-064-337	Sequence 337, App
	37	338	18.6	3278	1	US-07-778-890A-1	Sequence 1, Appli
	38	335	18.4	1230025	4	US-09-198-452A-1	Sequence 1, Appli
	39	335	18.4	1230230	4	US-09-438-185A-1	Sequence 1, Appli
	40	330	18.1	124884	4	US-09-661-596A-76	Sequence 76, Appli
	41	330	18.1	124884	4	US-09-913-514-1	Sequence 1, Appli
	42	330	18.1	125157	4	US-09-913-514-2	Sequence 2, Appli
	43	317	17.4	3678	4	US-09-112-580-13	Sequence 13, Appli
	44	307	16.9	35100	2	US-08-770-379-19	Sequence 19, Appli
	45	307	16.9	35100	3	US-08-757-669A-19	Sequence 19, Appli

ALIGNMENTS

RESULT 1
US-10-019-733-2
; Sequence 2, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619W00P
; CURRENT APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-2

Alignment Scores:
Pred. No.: 6.06e-227 Length: 1053
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-698-228-1 (1-351) x US-10-019-733-2 (1-1053)

QY 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSer 20

Db	1	ATGGGGACCCGAAAGCGCGAGCGCGCGCTGGATCAGATGAGACATCATCTTCA	60
Qy	21	AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg	40
Db	61	GACACCAACGAAAGTGAATAAAGTCAATGAGAGCCACTCTCTAGAAAGAGATTCTCGC	120
Qy	41	ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln	60
Db	121	CGGTTTGTCATCTTTCCAAATCCAGTACCCCTGATATTGGAAAAATGTATAAACAGGCACAG	180
Qy	61	AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisIleTrpAsnLys	80
Db	181	GCCTTCCTTCGGACAGCAGAAGAGGTGCATCTATCAAGAGGATCTCCCTCACTCGAACCAAG	240
Qy	81	LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaIaSerAsp	100
Db	241	CTTAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTAGCCCTTTTTCGAGCCAGTGAT	300
Qy	101	GlyIleValAsnGluAsnLeuValGluArgPheSerSerGlnGluValGlnValProGluAla	120
Db	301	GGAAATGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCACAGAGCT	360
Qy	121	ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu	140
Db	361	CGCTGTGTTCTATGGCTTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTG	420
Qy	141	LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu	160
Db	421	CTGATGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTATTAATGCAATTGAA	480
Qy	161	ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys	180
Db	481	ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCCCTTGCATGCGATAGCAGATAGAAAA	540
Qy	181	SerThrPheGlyGluArgValAlaIaPheAlaAlaValGluGlyValPhePheSerGly	200
Db	541	TCVACTTTTGGGAAAGAGTGGTGGCCCTTTGCTGCTGTAGAAAGAGTFTTCTCTCAAGGA	600
Qy	201	SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer	220
Db	601	TCITTTTGCTGCTATATTCUGCTTAAGAGAGAGGTCCTTATGTCAGAGACTCACTTTTTC	660
Qy	221	AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln	240
Db	661	AATGAACCTCATCAGCAGATGAAGGACTTCACTGTGACTTGTGCTTGCCTGATGTTCCAA	720
Qy	241	TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys	260
Db	721	TACTTTAGTAATTAAGCCTTCAGAAAGAGGTCAGGGAGATCATGTTGATGCTGTCAAA	780
Qy	261	IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle	280
Db	781	ATTGAGCAGAGATTTTAAACAGAGCCTTGCCAGTTGGCCTCATTTGGAATGATTCGATT	840
Qy	281	LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer	300
Db	841	TTGATGAACAGTACATATGAGTTTGTAGCTGACAGATTACTTGTGGAACTTGGATTCCTCA	900
Qy	301	LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys	320
Db	901	AAGGTTTTTCAGGCAGAAAAATCCTTTTGTATTTATGAAAAACATTTCTTTAGAAAGGAANA	960
Qy	321	ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr	340
Db	961	ACAAATTTCTTGGAAACCGAGTTTCAAGATATCAGCGTTTTCGAGTTATGGCGAARACC	1020
Qy	341	ThrAspAsnValPheThrLeuAspAlaAspPhe	351
Db	1021	ACAGATAACGCTCTTCACCTTGGATGTCAGATTTT	1053

```

; Sequence 12, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR FILING DATE: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 12
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-12

Alignment Scores:
Pred. No.: 6.06e-227 Length: 1053
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

```


QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 601 TCTTTTGTCTGCTATATTCTGGCTAAAGAAAGAGAGGCTTATGCGAGACTCACTTTTTC 660
QY 221 AsnGluLeuLysLeuSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
DB 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleValAspAlaValLys 260
DB 721 TACTTAGTAAATAGCCCTTCAGAGAAAGGCTCAGGAGATCATTTGATGCTGTCAAA 780
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
DB 781 ATTGAGCAGGAGCTTTTAAACAGAGCCCTTGCAGCTTGGCTCATTTGGAATGAATTCATT 840
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
DB 841 TTGATGAACACAGTACATAGATTGAGTTGTAGCTGACAGATTACTTGTGGAACCTTGATTTCTCA 900
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
DB 901 AAGGTTTTTCAGCGAGAAATCCTTTTGATTTTATGGAACACATTTCTTTAGAGGAAAA 960
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
DB 961 ACNAATTTCTTTGAGAAAGAGTTTCAGAGATATCAGGTTTTCAGATTATGCGAAGAAC 1020
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB 1021 ACAGATAACGCTTTCACCTTGGATGACAGATTTT 1053

RESULT 3

US-10-019-733-4
; Sequence 4, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 4
; LENGTH: 1081
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-4

Alignment Scores:

Pred. No.: 6,32e-227 Length: 1081
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-698-228-1 (1-351) x US-10-019-733-4 (1-1081)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
DB 20 ATGGCGCGACCCGGAAGCGCGGAGCGCGCGCTGATCAGGATCAGAGATCATCTTCA 79
QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40

DB 80 GACACCAACGAAAGTGAATAAAGTCAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 139
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
DB 140 CGGTTTGTCTATCTTCCAAATCCAGTACCCCTGATATTTGGAAAAATGATATAAACAGCACAG 199
QY 61 AlaSerPheTrpThrAlaGluGluValAspLysSerLysAspLeuProHisTrpAsnLys 80
DB 200 GCTTCTCTTCTGGACAGCAGAGAGGTTGACTTATCAAGGATCTCCCTCACTCGGAACAAG 259
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAspSerAsp 100
DB 260 CTTAAAGCAGATCAGAAAGTACTTCTCATCTCACATCTTAGCCTTTTTCAGCCAGCATGAT 319
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 320 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGGCT 379
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 380 CGCTGTTTCTATGCTTTCAAATTCATCGAGATGTTCACTCAGAGATGTCAGTTTG 439
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 440 CTGATAGACACTTACATCAGATCCCAAGAAAGCGAATTTTATTTAATGCAATTGAA 499
QY 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
DB 500 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGGATGGATGAGATAGAAAA 559
QY 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
DB 560 TCTACTTTTGGGAAAGAGTGGGCTTCTGCTCTGTAGAGAGTTTCTTCTCAGA 619
QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 620 TCTTTTGTCTATATTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 679
QY 221 AsnGluLeuLysSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
DB 680 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 739
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
DB 740 TACTTAGTAAATAAGCCTTCAGAAAGAAAGGTCAGGAGATCATTTGATGCTGTCAAA 799
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
DB 800 ATTGAGCAGGAGTTTTCAGAGAGCCCTTGCAGTTCCTCATTTGGAATGAATGCAATT 859
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
DB 860 TTGATGAACACAGTACATAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTTCTCA 919
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
DB 920 AAGGTTTTTCAGCGAGAAATCCTTTTGATTTTATGGAACACATTTCTTTAGAGGAAAA 979
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
DB 980 ACNAATTTCTTTGAGAAAGCAGTTTCAGAGATCAGGTTTTCAGATTATGCGAAGAAC 1039
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB 1040 ACAGATAACGCTTTCACCTTGGATGACAGATTTT 1072

RESULT 4

US-10-019-733-3
; Sequence 3, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA

Db 1202 TTTTAGCAAGGTTTTCAGAGTAGAAGAACCCATTGTGACTTTATGGAGAATATTTCACTGGA 1261
Qy 318 uGlyLysThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetal 338
Db 1262 AGGAAAGACTAATCTTTTGAAGAGAGATAGCGGAGTATCAGAGGATGGAGTGATGTC 1321
Qy 338 aGluThrThrAspAsnValPheThrLeuAlaAspPhe 351
Db 1322 AAGTCCAACAGAGAATCTTTTACCTTGGATGCTGACTTC 1361

RESULT 9

US-09-962-677-9
; Sequence 9, Application US/09962677
; Patent No. 6759200
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: THYMIDINE PHOSPHORYLASE GENE SEQUENCE
; TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING
; TITLE OF INVENTION: THE TREATMENT OF DISEASE
; FILE REFERENCE: 11926-015003
; CURRENT APPLICATION NUMBER: US/09/962,677
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 128, 1464
; OTHER INFORMATION: n = g or a
; NAME/KEY: misc_feature
; LOCATION: 189
; OTHER INFORMATION: n = t or g
; NAME/KEY: misc_feature
; LOCATION: 524
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc_feature
; LOCATION: 1399
; OTHER INFORMATION: n = t or a
; NAME/KEY: misc_feature
; LOCATION: 1636, 1738, 2259
; OTHER INFORMATION: n = c or t
US-09-962-677-9

Alignment Scores:
Pred. No.: 2,09e-179 Length: 2500
Score: 1460.50 Matches: 277
Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
DB: 4 Gaps: 1

US-10-698-228-1 (1-351) x US-09-962-677-9 (1-2500)

Qy 4 ProGluArgProGluAlaGlyLeuAspGluAspGlu-ArgSerSerSerAspThrAs 23
Db 302 CTGACGGAGACCCCGCTCTGCGCCAGCAAGACCGGAGGAGGATCTTCCAGGAGCCAC 361
Qy 23 nGluSerGluLeuLysSer-----AsnGluGluProLeuLeuArgLysSe 38
Db 362 GGAGCCGAAACTAAAGACGACTGCCCGCGGTGGAGGATGAGCCGCTGCTGAGAGAAA 421

Qy 38 rSerArgAtqPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGl 58
Db 422 CCCCCCGCGTCTTGTCACTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481
Qy 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
Db 482 GGCAGAGGCTTCTTTTGGACCGCCGAGGAGGTTGACCTCTCNAAGGACATTCAGCACTG 541
Qy 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAl 98
Db 542 GGAATCCCTGAAACCCGAGGAGATATTTATATCCATCTTCTGGCTTTCTTTTGCAGC 601
Qy 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
Db 602 AAGCATGGCATAGTAATAAGAAACTTGGTGGAGCGATTATAGCCAAAGATTCAGATTAC 661
Qy 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
Db 662 AGAAGCCCGCTGTTCTATGCTTCCAAATTTGCCATGGAAACATACATTCCTGAAATGTA 721
Qy 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
Db 722 TAGTCTTCTTATTGACACTTACATAAAAGATCCCAAGAAAGAGGAAATTTCTTCAATGC 781
Qy 158 aIleGluThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTTrpIleAlaAs 178
Db 782 CATTTGAAACCATGCTTGTGTCAAGAAAGAGCGAGACTGGGCGCTTGGCCCTGATTTGGGA 841
Qy 178 pArgLysSerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePh 198
Db 842 CAAAGAGGCTACCTATGTTGACGTTGTAGCCCTTTGCTCAGTGGAGGACATTTCTT 901
Qy 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
Db 902 TTCGGTCTTTTGGCGTCGATATTTCTGCTCAAGAAACGAGGACTGATGCTCGGCTCAC 961
Qy 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
Db 962 ATTTTCTAATGAACCTTATTAGCAGATGAGGGTTTACACTGTGATTTTGTCTGCTGAT 1021
Qy 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
Db 1022 GTTCAACACCTGGTACACAAACCATCGAGGAGAGAGTAAGAGAAATAATTATCAATGC 1081
Qy 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMeAs 278
Db 1082 TGTTCGGATAGAACAGAGTTCCTCACTGAGGCTTGCCTCTGAAGCTCAITGGGATGAA 1141
Qy 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuGl 298
Db 1142 TTGCACCTCTAATGAAGCAATACATTTAGTTTGTGGCAGACAGACTTATGCTGGAACTGG 1201
Qy 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGl 318
Db 1202 TTTTAGCAAGTTCCTCAGAGTAGAAGAACCACTTTGACTTTATGGAGATATTTCACTGGA 1261
Qy 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMeAl 338
Db 1262 AGGAAAGACTAATCTTTTGAAGAGAGATAGGCGAGTATCAGAGGATGGGAGTGATGTC 1321
Qy 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1322 AAGTCCAACAGAGAATCTTTTACCTTGGATGCTGACTTC 1361

RESULT 10

US-09-949-016-145
; Sequence 145, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

```

; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 145
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-145

Alignment Scores:
Pred. No.:      2,09e-179      Length:      2500
Score:          1460.50      Matches:      277
Percent Similarity: 88.42%      Conservative: 36
Best Local Similarity: 78.25%      Mismatches: 35
Query Match:      80.20%      Indels:      6
DB:               4          Gaps:      1

US-10-698-228-1 (1-351) x US-09-949-016-145 (1-2500)

QY 4 ProGluArgProGluAlaAlaGlyLeuAspGluArgSerSerSerAspThrAs 23
DB 302 CCGTGGCGGAGCCCGCGTCTCGCCAGCAGCAAGACCGCGAGGAGGATCTCCAGGAGCCAC 361
QY 23 nGluSerGluLeuLysSer-----AsnGluGluProLeuLeuArgLysSe 38
DB 362 GGAGCGCGAAATAAGCAGAGCTGCCCGCGGTGGAGGATGAGCGCGCTCTGAGAGAAA 421
QY 38 rSerArgArgPheValIlePheProIleGlnTyProAspIleTrpLysMetTyLysG 58
DB 422 CCCCAGCGCGCTTGTCACTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481
QY 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAlaAspLeuSerLysAspLeuProHisTr 78
DB 482 GGCAGAGGCTCTCTTTGGACCGCGCGAGGAGGTGACCTCTCCAAGGACATTCAGCACTG 541
QY 78 pAsnLysLeuLysAlaAspGluLysTyPheIleSerHisIleLeuAlaPhePheAlaAl 98
DB 542 GGAATCCCTGAAACCGAGGAGAGATATTTATATCCCATGTCTGGCTTTCTTTCAGC 601
QY 98 aSerAspGlyIleValAsnGluAlaValGluArgPheSerGlnGluValGlnValPr 118
DB 602 AAGCATGGCATAGTAAATGAAACCTTGGTGAGCGATTTAGCCAGAAAGTTCAAGATTAC 661
QY 118 oGluAlaArgCysPheTyGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
DB 662 AGAAGCCCGCTGTTCTATGCTTCCAAATTGCCATGAAACATACATCTCTGAAATGTA 721
QY 138 rSerLeuLeuIleAspThrTyIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
DB 722 TAGTCTTCTTATTCACACTTACATAAAGATCCCAAGAAAGGGAATTTCTCTTCAATGC 781
QY 158 aIleGluThrMetProTyValLysLysLysValaAspTrpAlaLeuArgTrpIleAlaAs 178
DB 782 CATTAAGACATGCTTGTGTCAAGAAAGGAGCAGACTGGGCTTGGCTGGATTTGGGA 841
QY 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePh 198
DB 842 CAAAGAGGCTACCTATGGTGAACGGTGTGATGCTTGTGCTGAGTGAAGGATTTCTT 901
QY 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
DB 902 TTCCGGTCTCTTGTGGTGCATATTTGGCTCAAGAAACGAGGAGCTGATCGCTGGCCCTAC 961
QY 218 rPheSerAnGluLeuIleSerArgAspGluLysLeuHisCysAspPheAlaCysLeuMe 238

```

```

DB 962 ATTTTCTAATGAACCTTATTAGCAGAGATGAGGGTTTACACCTGTGATTTTGTGCTGAT 1021
QY 238 tPheGlnTyLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
DB 1022 GTTCAACACCTGTGTACACAAACCATCGAGGAGAGAGTAAGAGAAATAATATCAATGC 1081
QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
DB 1082 TGTTCGATAGAACAGAGGATTCCTCACTGAGGCCCTTGCCTGTGAAGCTCAITGGGATGA 1141
QY 278 nCysIleLeuMetLysGlnTyIleGluPheValAlaAspArgLeuValGluLeuGl 298
DB 1142 TTGCACCTTAATGAAGCAATACATTTGATTTGTGCGACAGACACTTATGCTGGAACCTGG 1201
QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGl 318
DB 1202 TTTTAGCAAGGTTTTCAGAGTAGAGAACCCCATTTGACTTTTATGGAGATAATTTTCACTGA 1261
QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGlnArgPheAlaValMetAl 338
DB 1262 AGGAAGACCTACTTCTTTGAGAAGACAGATAGGCGAGTATCAGAGGATGGGAGTGTATC 1321
QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB 1322 AAGTCCAACAGAGAAATTTCTTTTACCTTGGATGCTGACTTC 1361

RESULT 11
US-08-307-499-1/c
; Sequence 1, Application US/08307499
; Patent No. 5651972
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; TITLE OF INVENTION: Live Vaccine Vector
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESS: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: U.S.A.
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/307,499
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,241
; FILING DATE: 1-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,630
; FILING DATE: 29-JUN-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/342,212
; FILING DATE: 21-APR-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF35.1.FWCCI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:

```

LENGTH: 14176 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:

NAME/KEY: CDS

LOCATION: 3852..4226

FEATURE:

NAME/KEY: CDS

LOCATION: 4585..4887

FEATURE:

NAME/KEY: CDS

LOCATION: 5131..5310

FEATURE:

NAME/KEY: CDS

LOCATION: 5760..5912

FEATURE:

NAME/KEY: CDS

LOCATION: 6786..7130

FEATURE:

NAME/KEY: CDS

LOCATION: 10148..10513

US-08-307-499-1

Alignment Scores:

Pred. No.: 1,37e-145 Length: 14176
Score: 1210.00 Matches: 233
Percent Similarity: 82.01% Conservative: 36
Best Local Similarity: 71.04% Mismatches: 58
Query Match: 66.45% Indels: 2
DB: 1 Gaps: 1

US-10-698-228-1 (1-351) x US-08-307-499-1 (1-14176)

QY 24 GluSerGluIleuSerAsnGluGluProLeuLeuArgLysSerArgArgPheVal 43
DB 10842 AAAAATGATTTTATACAGAAAGTGGAGCTATCTTCAGAGTCTGATCTAGTTCGTT 10783
QY 44 IlePheProIleGlnTyrProIleTyrLysMetTyrLysGlnAlaGlnAlaSerPhe 63
DB 10782 ATTTCCCTATTAAAGTATCATGATATCTCGAAATGTATAAACAATCAGTGGCAAGTTT 10723
QY 64 ThrThrAlaGluGluValAspLysSerLysAspLeuProHisTyrAsnLysLeuAla 83
DB 10722 TGGACCGTTGAGAAAGTATGATTTATCAAAAGATTTAGATGATTGGGATAAATTAACATA 10663
QY 84 AspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAspGluIleVal 103
DB 10662 GACGAAAATATCTTTATAAACAATATAGCAATTTTTCATCTAGTATGATTTGTA 10603
QY 104 AsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAlaArgCysPhe 123
DB 10602 AATGAAATTTAGCGGAAAGATTTTATGTTGGATGTACAGTGTTCAGAGGCACGATGTTTC 10543
QY 124 TyrGlyPheGlnIleLeuLeuGluAsnValHisSerGluMetTyrSerLeuLeuIleAsp 143
DB 10542 TATGGATTTCAATAGCTATGGAATAATATTCATTCAGAAATGTATGATTTTATTAATAGAT 10483
QY 144 ThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGluThrMetPro 163
DB 10482 ACATATGTAAGAGATATATAGAAAAATGTCATTTTAACTATAGAACAAATGGAA 10423
QY 164 TyrValLysLysLysAlaSerTrpAlaLeuArgTrpIleAlaAspArgLysSerThrPhe 183
DB 10422 TCGTAAAAAAGAAAGCTGATTGGGCCAGAAATGATATCTAGCAACAAAG---GTATAT 10366
QY 184 GlyGluArgValValAlaPheAlaAlaValGluGluValPhePheSerGlySerPheAla 203
DB 10365 GGAGAAAGAGTAGTAGCAATTTGACAGCTGTGGAGGAATATCTTTCTGCTTCTATTTGCT 10306
QY 204 AlaIlePheThrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeu 223

DB 10305 GCTATATTTTGGATAAAAAACGAGGATTCATCCCGGATTAACATTTCTTAATGAACATA 10246
QY 224 IleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGlnTyrLeuVal 243
DB 10245 ATAAGTAGAGACGAAAGGTTTACATTTGTGATTTTCGCGTGTAAATGTTTAAACATTTATTA 10186
QY 244 AsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLysIleGluGln 263
DB 10185 CATCCACCATCTAAGGAAGTTATAACGTCGATATCATTCGCGTGTAAATGTTTAAACATTTATTA 10126
QY 264 GluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIleLeuMetLys 283
DB 10125 GAGTTTTTGACAGTTGCTATTCGCGTGTATATAGTATGATTTGTTTAAATGTTCT 10066
QY 284 GlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSerLysValPhe 303
DB 10065 CAGTATATAGAAATTCGTCGAGATAGATTTATTAACAGAGTTAGTTGTGAAAAGT--CTC 10008
QY 304 GlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLysThrAsnPhe 323
DB 10007 AATGATATATCTCTTTTATGAGTATATATCATCTAGAGGTTAAGACTAATTTT 9948
QY 324 PheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThrThrAspAsn 343
DB 9947 TTCGAACGACGAGTTAGTGAATATCAAAAGATGGCGGTGTTTACAAATAAAGAGAGAAT 9888
QY 344 ValPheThrLeuAspAlaAspPhe 351
DB 9887 ATATTTCTACGGATATAGATTTT 9864
RESULT 12
US-08-307-499-14
; Sequence 14, Application US/08307499
; Patent No. 5651972
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: U.S.A.
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/307,499
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,241
; FILING DATE: 1-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,630
; FILING DATE: 29-JUN-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/342,212
; FILING DATE: 21-APR-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF35.1.FWCCI
; TELECOMMUNICATION INFORMATION:

TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 14176 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:

NAME/KEY: CNS

NAME/ID:	CDS
LOCATION:	139 1450

NOTATION: 00471..T460
FEELING.

FEATURE: NAME / KEY: CNO

NAME/KEY:	CDS
LOCATION:	REF

LOCATION: 2456..2659

FEATURE:

NAME/KEY: CDS

LOCATION: 2809..3030

FEATURE:

NAME/KEY: CDS

LOCATION: 3070..3330

FEATURE:

NAME/KEY: CDS

LOCATION: 3356..4180

FEATURE:

NAME/KEY: CDS

LOCATION: 4392..5894

FEATURE:

NAME/KEY: CDS

LOCATION:	6171	6398
-----------	------	------

LOCATION: 01/11.0530
FEATURE:

NAME / REV. : CNO

NAME/KEY:	CDS
LOCATION:	C117 C07E

LOCATION: 044/..08/5
PRATT.

FEATURE: **1997/1998** **2000**

NAME/KEY:	CDS
FACTS	000

LOCATION: 6928..7431

FEATURE:

NAME/KEY: CDS

LOCATION: 7454..7858

FEATURE:

NAME/KEY: CDS

LOCATION: 7895..8155

FEATURE:

NAME/KEY: CDS

LOCATION: 8215..8682

Alignment Scores:

[illegible]

RESULT 13

US-09-299-268-1/c

US-09-299-208-17C
: Sequence 1, Application US/09299268

Patent No. 6217882
 GENERAL INFORMATION:
 APPLICANT: Moyer, Richard W.
 APPLICANT: Vi uela, Eladio
 APPLICANT: Gibbs, E.P.J.
 TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
 TITLE OF INVENTION: Live Vaccine Vector
 NUMBER OF SEQUENCES: 60
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: David R. Saliwanchik
 STREET: 2421 N.W. 41st Street, Suite A-1
 CITY: Gainesville
 STATE: Florida
 COUNTRY: U.S.A.
 ZIP: 32606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/299,268
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/901,127
 FILING DATE:
 APPLICATION NUMBER: US 07/908,241
 FILING DATE: 1-JUL-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/908,630
 FILING DATE: 29-JUN-1992
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/342,212
 FILING DATE: 21-APR-1992
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Saliwanchik, David R.
 REGISTRATION NUMBER: 31,794
 REFERENCE/DOCKET NUMBER: UF35.1.FWCC1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 904-375-8100
 TELEFAX: 904-372-5800
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 14176 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: unknown
 MOLECULE TYPE: DNA (genomic)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 3852..4226
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 4585..4887
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 5131..5310
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 5760..5912
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 6786..7130
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 10148..10513
 US-09-299-268-1
 Alignment Scores:
 Pred. No.: 1.37e-145 Length: 14176

Score: 1210.00 Matches: 233
 Percent Similarity: 82.01% Conservative: 36
 Best Local Similarity: 71.04% Mismatches: 58
 Query Match: 66.45% Indels: 2
 DB: 3 Gaps: 1
 US-10-698-228-1 (1-351) x US-09-299-268-1 (1-14176)
 QY 24 GluSerGluLeuLeuLysSerAsnGluGluProLeuArgLysSerSerArgPheVal 43
 DB 10842 AAAATGATTTTATACGAATGGAGCCCTATCTTCAAGAGCTGATTTCTAGGTTCGTT 10783
 QY 44 IlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGlnAlaSerPhe 63
 DB 10782 ATTTCCCTATTAAAGTATCATGATATCTGAAAATGTATAACAATCAGTGGCAAGTTT 10723
 QY 64 TrpThrAlaGluValAspLeuSerLysAspLeuProHisTyrAsnLysLeuLysAla 83
 DB 10722 TGGACCGTTTGAAGAAGTAGATTTATCAAAAGATTAGATGATTTGGGATAAATTAACATAA 10663
 QY 84 AspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAspGlyIleVal 103
 DB 10662 GACGAAAAATACTTTATAAACAATATCTAGCATTTTTCGATCTAGTATGGTATTTGTA 10603
 QY 104 AsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAlaAlaArgCysPhe 123
 DB 10602 AATGACAATTTAGCGGAAGATTTTATGTGATGTACAGTGTTCAGAGGCACGATGTTTC 10543
 QY 124 TyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeuLeuIleAsp 143
 DB 10542 TATGGATTTCAAAATAGCTATGGAAAATATTCATTCAGAAAATGTATAGTTTATTATAGAT 10483
 QY 144 ThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGluThrMetPro 163
 DB 10482 ACATATGTAGAGATAATATAGAAAAAATGCAATTTATTTCACGCTATAGAACCAATGAA 10423
 QY 164 TyrValLysLysLysAlaAspTyrAlaLeuArgTyrPheAlaAspArgLysSerThrPhe 183
 DB 10422 TCGTAAAAAAGAAAGCTGATTGGCCAGAAAATGGATATCTAGCAACAAAG--GTATAT 10366
 QY 184 GlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGlySerPheAla 203
 DB 10365 CGAAGAAGAGTAGTAGCATTTGCAGCTGTGAGGGAATATTTCTTTCTGTTCTCATTTGCT 10306
 QY 204 AlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeu 223
 DB 10305 GCTATATTTTGGATAAAAAAACGAGGATGATGCCCGGATTAACATTTTCTTAATGAACATA 10246
 QY 224 IleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGlnTyrLeuVal 243
 DB 10245 ATAAAGTAGAGCAAGGTTTACATTTGTGATTTTGGGTGTTTAAATGTTTAAACATTTATTA 10186
 QY 244 AsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLysIleGluGln 263
 DB 10185 CATCCACCCTCTAAGGAAGTTATAACGTCGATATCATTCATGCGGTAAATATAGAAAAG 10126
 QY 264 GluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIleLeuMetLys 283
 DB 10125 GAGTTTTTGACAGTTGCTATCCGGTGGATCTTATAGGTATGAATTTGTTTAAATGCT 10066
 QY 284 GlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSerLysValPhe 303
 DB 10065 CAGTATATAGAAATTCGTCGAGATAGATTTAATACAGAGTAGGTAGTTGTGAAAAGT--CTC 10008
 QY 304 GlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLysThrAsnPhe 323
 DB 10007 AATGTATATAATCCTTTTACGTTTATCGAGTATATATCACTAGAAAGGTAAAGCTAATTTT 9948
 QY 324 PheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThrThrAspAsn 343
 DB 9947 TTCGAACGACGAGTTAGTGAATATCAAAAGATGGGGGTGTTTACAAAATAAAGAGAGAT 9888
 QY 344 ValPheThrLeuAspAlaAspPhe 351

```

Db      9887 ATATTCTACGGATAGATATT 9864
      :...|||::: ||| |||||
RESULT 14
US-09-299-268-14
; Sequence 14, Application US/09299268
; Patent No. 6217882
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSER: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: U.S.A.
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/299,268
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/901,127
; FILING DATE:
; APPLICATION NUMBER: US 07/908,241
; FILING DATE: 1-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,630
; FILING DATE: 29-JUN-1992
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/342,212
; FILING DATE: 21-APR-1992
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF35.1.FWCC1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14176 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 138..1460
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2456..2659
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2809..3030
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3070..3330
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3356..4180
; FEATURE:

; NAME/KEY: CDS
; LOCATION: 4392..5894
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 6171..6398
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 6447..6875
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 6928..7431
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 7454..7858
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 7895..8155
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 8215..8682
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 8715..9539
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 9562..10272
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 10316..11908
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 11971..12780
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 12829..13107
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 13149..14171
; US-09-299-268-14

Alignment Scores:
Pred. No.: 1.37e-145 Length: 14176
Score: 1210.00 Matches: 233
Percent Similarity: 82.01% Conservativity: 36
Best Local Similarity: 71.04% Mismatches: 58
Query Match: 66.45% Indels: 2
DB: 3 Gaps: 1

US-10-698-228-1 (1-351) x US-09-299-268-14 (1-14176)
QY 24 GluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArgArgPheVal 43
Db 3335 AAAAATGATTTTATACACGAATGGAGCTATTCTTCAAGAGTCTGATTCTAGGTTGTT 3394
QY 44 IlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGlnAlaSerPhe 63
Db 3395 ATTTTCCCTATTAAAGTATCATGATATCTGAAAATGTATAAACAATCAGTGCAGATTTT 3454
QY 64 TrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLysLeuLysAla 83
Db 3455 TGGACCGTTGAAGAAGTAGATTATCAAAAGATTTAGATGATGGATAAATTAACTAAA 3514
QY 84 AspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaIaSerAspGlyIleVal 103
Db 3515 GACGAAAAATACTTTATAAAACATATACTAGCATTTTTCATCTAGTATGTTATGTA 3574
QY 104 AsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAlaArgCysPhe 123
Db 3575 AATGAGAAATTTACGGAAAGATTTTATGTGTGTGTACAGTGTTCAGAGGCACGATGTTTC 3634
QY 124 TyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeuLeuIleAsp 143
Db 3635 TATGGATTTCAAAATAGCTATGGAAATATTCAATTCAGAAATGATAGTTTATTATAGAT 3694
```

```
QY 144 ThrTyrIleArgProLysLysArgGluPheLeuPheAsnAlaIleGluThrMetPro 163
Db 3695 ACATAATGAAGAGATAATATAGAAAATATGCATTATTTAACCTATAGAACATGGAA 3754
QY 164 TyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLysSerThrPhe 183
Db 3755 TCGGTAAAAAGAAAGCTGATTGGGCCAGAAAATGATATCTAGCAACAAG--GTATAT 3811
QY 184 GlyGluArgValAlaPheAlaValGluGlyValPhePheSerGlySerPheAla 203
Db 3812 CGAAGAAGAGTAGTAGCATTTTCAGCTGTGGAGGGAATATCTTTCTCGTTCATTGCT 3871
QY 204 AlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeu 223
Db 3872 GCTATATTTGGATATAAAAACAGAGATTGATGCCCGGATTACATTTCTATATGAACTA 3931
QY 224 IleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGlnTyrLeuVal 243
Db 3932 ATAAGTAGAGACGAGGTTTACATTGTGATTTTGGCTGTTTAAATGTTTAAACATTTATTA 3991
QY 244 AsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLysLysGluGln 263
Db 3992 CATCCACCATTAAAGGAAGTTTAAACGTCGATAATCATTTGATGCGGTTTAAATAGAAAAG 4051
QY 264 GluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIleLeuMetLys 283
Db 4052 GAGTTTTTACAGAGTGTCTATTCGGTGGATCTTATAGGTATGATGTTTGTGTTTATGTCT 4111
QY 284 GlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSerLysValPhe 303
Db 4112 CAGTATATAGATTCGTCGACATAGATATTTAAACAGAGTAGTTGTGTAAGT--CTC 4169
QY 304 GlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLysThrAsnPhe 323
Db 4170 AATGTATATAATCCTTTTAGCTTTATGGAGTATATATCACTAGAGGTAAAGACTAATTTT 4229
QY 324 PheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThrThrAspAsn 343
Db 4230 TTCGAACGACGAGTTAGTGAATATCAAAAGATGGGGGTGTTTACAAATAAAGAAGAAT 4289
QY 344 ValPheThrLeuAspAlaAspPhe 351
Db 4290 ATATTTTCTACGATATAGATTTT 4313

RESULT 15
US-09-248-796A-3914
; Sequence 3914, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 3914
; LENGTH: 1236
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-3914

Alignment Scores:
Pred. No.: 5,45e-137 Length: 1236
Score: 1131.00 Matches: 210
Percent Similarity: 78.21% Conservative: 52
Best Local Similarity: 62.69% Mismatches: 67
Query Match: 62.11% Indels: 6
```

```
DB: 4 Gaps: 3
US-10-698-228-1 (1-351) x US-09-248-796A-3914 (1-1236)
QY 23 AsnGluSerGluIleLysSerAsnGlu-----GluProLeuLeuArgLysSerSerArg 40
Db 229 AACAGACACAAAGTTAAACACAGAGAGTCCCAACAGAGCAATTGCTTACACACAGACAAGACG 288
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
Db 289 CGTCATACCATATATCGATAAAAGATATCCCGAGTTTGGCAGTTCTATAAAAGAGTCGTTA 348
QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrp---Asn 79
Db 349 GCCTCATTTCTGGACTGCTGAAGAGCTTGACTTGAGTAAAGATTGGATGATGGGAATAAT 408
QY 80 LysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSer 99
Db 409 AAGATGAACCAACCAACGAAAGATTCTTTATCTCTCGAGTTTGGCATTTCTTTGCTGTCATCG 468
QY 100 AspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGlu 119
Db 469 GATGGTATTGTTAATGAGAACTTGGTGGAAATTTCTGTCAGAGTGCACAAATACCCGAG 528
QY 120 AlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSer 139
Db 529 GCTAAACTGGTTTACAAGTTTCAGATCATGTCGAGAAACATTCATTCAGAGACTTATTCG 588
QY 140 LeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIle 159
Db 589 TTGTTAATAGAAACATATTTCAAAGACCCCGAAGACAGACTTTTGTGTTAATGCTATT 648
QY 160 GluThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArg 179
Db 649 GACAATATCCCATTTATAAGAAAAAAGCCGATTTGGGCAATTCAGATGGATCCAGACGAG 708
QY 180 LysSerThrPheGlyGluArgValAlaPheAlaValGluGluValPhePheSer 199
Db 709 GATGCTTTGTATGCTGAGAGATTGGTTGCACTTCTGCTGTTGAAGGTATATTTTTCAGT 768
QY 200 GlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPhe 219
Db 769 GGCTCGTTTGCAGCAATTTCTGTTTAAAGAAACCGGGTTTGTATGCCGGGTGACATTT 828
QY 220 SerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPhe 239
Db 829 TCCAATGAATTAATTTTGCAGACGAAAGGGATTCCACACAGATTTATGCCCTGTTTATTTT 888
QY 240 GlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaVal 259
Db 889 TCATATTTTAAAAATAAACCTTCTTCAGAAATAATTGAGAAAAATCATCCAGAGCTGTT 948
QY 260 LysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCys 279
Db 949 GATATCGAAGAAAGAAATATTTTCAGCGATCGCTTCCAGTATCTTATTGGGAATGAATGT 1008
QY 280 IleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPhe 299
Db 1009 GACTTCGATGTGCCAGTATGTTGAATTTGTTGCTGATAGATTACTAGTAGCTTTTGGAAAT 1068
QY 300 SerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGly 319
Db 1069 AAAAGTACTACAAATGTTTACCAACCACTTTGATTTTATGGAATAATTTCTTGGCTGCT 1128
QY 320 LysThrAsnPhePheGluLysArgValSerGlnTyrGlnArgPheAlaValMetAlaGlu 339
Db 1129 AAACTAATTTCTTTGAAAAAGAGAGTCTCTGATTATCAAAAGGCTGGGTGTGTATGGA 1188
QY 340 Thr-----ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1189 GTGGAAAAATAAGAATGAGAAAACTGGATTGTTTGTTCATCAAGATTTC 1233
```

Search completed: October 30, 2005, 08:56:03
Job time : 185 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 30, 2005, 08:09:55 ; Search time 661 Seconds
(without alignments)
4385.636 Million cell updates/sec

Title: US-10-698-228-1
Perfect score: 1821
Sequence: 1 MGDPERPBAAGLDQDERSSS.....QRFAVMAETTDNVFTLDADF 351

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9784742 seqs, 4129495052 residues

Total number of hits satisfying chosen parameters: 19569484

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -MODE=frame+pn.model -DEV=xlh
-Q=/cgm2_1/USPTO_spool/US10698228/runat_26102005_100716_5225/app_query.fasta_1.519
-DB=Published Applications NA -OPMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40 cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100
-MINLEN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10698228 @CGN_1_1_354 @runat_26102005_100716_5225
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgm2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgm2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
3: /cgm2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
4: /cgm2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgm2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
6: /cgm2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
7: /cgm2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
8: /cgm2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgm2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:
10: /cgm2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:
11: /cgm2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:
12: /cgm2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:
13: /cgm2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:
14: /cgm2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:
15: /cgm2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:
16: /cgm2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:
17: /cgm2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:
18: /cgm2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:
19: /cgm2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:
20: /cgm2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:
21: /cgm2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:
22: /cgm2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:
23: /cgm2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:
24: /cgm2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:
25: /cgm2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:
26: /cgm2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:
27: /cgm2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:
28: /cgm2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1821	100.0	1053	19	US-10-698-228-2
2	1821	100.0	1053	19	US-10-698-228-12
3	1821	100.0	1081	19	US-10-698-228-4
4	1821	100.0	4955	18	US-10-172-118-71
5	1821	100.0	4955	19	US-10-342-887-71
6	1821	100.0	4955	19	US-10-698-228-3
7	1460.5	80.2	1989	9	US-09-925-301-505
8	1460.5	80.2	2216	16	US-10-084-817-342
9	1460.5	80.2	2482	24	US-10-220-335-514
10	1460.5	80.2	2500	9	US-09-954-456-724
11	1460.5	80.2	2500	9	US-09-954-456-1169
12	1460.5	80.2	2500	9	US-09-954-456-1827
13	1460.5	80.2	2500	19	US-10-641-643-1370
14	1460.5	80.2	2500	21	US-10-733-878-458
15	1460.5	80.2	2500	22	US-10-843-641A-3751
16	1460.5	80.2	2500	22	US-10-843-641A-4196
17	1460.5	80.2	2500	22	US-10-843-641A-4854
18	1460.5	80.2	2500	24	US-10-756-149-713
19	1447.5	79.2	2641	24	US-10-220-335-170
20	1443	79.2	2113	22	US-10-764-420-1636
21	1425.5	78.3	1328	19	US-10-403-571-75
22	1376.5	75.6	1371	24	US-10-450-763-15278
23	1348	74.0	1289	26	US-11-097-143-19181
24	1280	70.3	186854	22	US-10-872-156-34
25	1241.5	68.2	3945	26	US-11-097-143-19180
26	1230	67.5	1218	16	US-10-128-714-7245
27	1220.5	67.0	1292	18	US-10-369-493-27809
28	1192.5	65.5	1146	16	US-10-128-714-2245
29	1192	65.5	1242	17	US-10-032-585-6537
30	1176.5	64.6	963	18	US-10-369-493-34052
31	1168.5	64.2	1200	18	US-10-369-493-45857
32	1161.5	63.8	1173	18	US-10-369-493-46252
33	1161.5	63.8	1206	18	US-10-320-797-2145
34	1151	63.2	1314	16	US-10-128-714-1245
35	1151	63.2	1450	16	US-10-128-714-6245
36	1151	63.2	3314	16	US-10-128-714-245
37	1151	63.2	3450	16	US-10-128-714-5245
38	1147.5	63.0	1586	19	US-10-424-599-73024
39	1133.5	62.2	1169	19	US-10-425-114-31361
40	1133.5	62.2	2992	21	US-10-425-115-31182
41	1132.5	62.2	1725	20	US-10-437-963-11109
42	1128.5	62.0	1255	20	US-10-767-701-13723
43	1125.5	61.8	1791	21	US-10-425-115-63321
44	1114.5	61.2	1248	19	US-10-424-599-126843
45	1113	61.1	1146	9	US-09-822-830A-174

ALIGNMENTS

RESULT 1
US-10-698-228-2
; Sequence 2, Application US/10698228
; Publication No. US2004007253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; PRIOR FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28

```
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-2
Alignment Scores:
Pred. No.: 1.55e-216 Length: 1053
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 19 Gaps: 0
US-10-698-228-1 (1-351) x US-10-698-228-2 (1-1053)
QY 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSer 20
DB 1 ATGGGCGACCCGGGAAGCGCGGAGCGCGGCTGGATCAGATCAGATCATCTTCA 60
QY 21 AspThrAsnGluSerGlnLeuLysSerAsnGluGluProLeuLeuArgLysSerArg 40
DB 61 GACACCAACGAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
DB 121 CGGTTTGTCTATCTTCCAAATCCAGTACCTGATATTTGGAAATGTATTAACAGGACAG 180
QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
DB 181 GCTTCTCTTGACAGCAGAGAGGTGCTGATATCAAGATCTCCCTACTGGAACAG 240
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp 100
DB 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTACATCTTAGCCTTTTTCAGGCGAGTAT 300
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnProGluAla 120
DB 301 GGAATTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 360
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 361 CGCTGTTTCTATGGCTTTTCAAAATCTCATCGAAGATGTTCACTCAGAGATGACAGTTTG 420
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 421 CTGATAGACACTTACATCAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
QY 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
DB 481 ACCATGCCCTATGTTAAGAAAAAGAGATTTGGGCTTGGATGGATAGCAGATAGAAAA 540
QY 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGluValPhePheSerGly 200
DB 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAAGAGAGTGTTCCTCTCAGGA 600
QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 601 TCTTTTGTCTATATCTTGGCTAAGAGAGAGAGGTCTTATGTCAGGACTCATTCTTCC 660
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
DB 661 AATGAACATCATCAGCAGATGAAGGACTTCACTGTGACTTTGCTGCTGCTGATGTTCCAA 720
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
; TACTTAGTAAATAAGCCTTCAGAAGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAAA 780
; IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
; ATTGAGCAGGAGTATTTAAACAGAGGCTTCCAGTTCGCTCATTTGGAAATGAATTCGATT 840
; LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
; TTGATGAAACAGTACATTTAGTGTGAGCTGACAGATTTACTTTGGAACTTTGGATCTCA 900
; LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
; AAGGTTTTTCAGGCAGAAAAATCCTTTTGAITTTATGGAACAATTTCTTTTAGAAGAAAA 960
; ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
; ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTTGCAGTTTATGCGAGAAAC 1020
; ThrAspAsnValPheThrLeuAspAlaAspPhe 351
; 1021 ACAGATAACGCTCTTCACCTTGGATGCAGATTTT 1053
RESULT 2
US-10-698-228-12
; Sequence 12, Application US/10698228
; Publication No. US20040072253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619W00P
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 12
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-12
Alignment Scores:
Pred. No.: 1.55e-216 Length: 1053
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 19 Gaps: 0
US-10-698-228-1 (1-351) x US-10-698-228-12 (1-1053)
QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
DB 1 ATGGGCGACCCGGGAAGCGCGGAGCGCGGCTGGATCAGATCAGATCATCTTCA 60
QY 21 AspThrAsnGluSerGlnLeuLysSerAsnGluGluProLeuLeuArgLysSerArg 40
DB 61 GACACCAACGAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
DB 121 CGGTTTGTCTATCTTCCAAATCCAGTACCTGATATTTGGAAATGTATTAACAGGACAG 180
QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
DB 181 GCTTCTCTTGACAGCAGAGAGGTGCTGATATCAAGATCTCCCTACTGGAACAG 240
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp 100
DB 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTACATCTTAGCCTTTTTCAGGCGAGTAT 300
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnProGluAla 120
DB 301 GGAATTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 360
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 361 CGCTGTTTCTATGGCTTTTCAAAATCTCATCGAAGATGTTCACTCAGAGATGACAGTTTG 420
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 421 CTGATAGACACTTACATCAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
QY 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
DB 481 ACCATGCCCTATGTTAAGAAAAAGAGATTTGGGCTTGGATGGATAGCAGATAGAAAA 540
QY 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGluValPhePheSerGly 200
DB 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAAGAGAGTGTTCCTCTCAGGA 600
QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 601 TCTTTTGTCTATATCTTGGCTAAGAGAGAGAGGTCTTATGTCAGGACTCATTCTTCC 660
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
DB 661 AATGAACATCATCAGCAGATGAAGGACTTCACTGTGACTTTGCTGCTGCTGATGTTCCAA 720
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
```

Db 181 GCTTCCTCTTGGACACAGAGAGTTGACTTATCAAGAGTCTCCCTCAGTGAACAAG 240
Qy 81 LeuLeuAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Db 241 CTTAAGCAGATGAGAGTACTTCTCTCATCTTACCTTTTGGAGCAGTGAT 300
Qy 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 301 CGAATTGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCCAGAGCT. 360
Qy 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 361 CGCTGTTCTATGGCTTTCAAAATCTCATGAGAATGTTTCACACAGATGTACAGTTG 420
Qy 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 421 CTGATAGACATTCATACAGATCCAGATCCAGAAAGGGAATTTTATTTAATGCAATGAA 480
Qy 161 ThrMetProTyrValLysLysLysAlaAspTyrPalaLeuArgTyrIleAlaAspArgLys 180
Db 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTGGCCCTTGCATGGATAGCAGATAGAAA 540
Qy 181 SerThrPheGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
Db 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAGAGTTCCTTCTCAGGA 600
Qy 201 SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 601 TCTTTTGGCTGCTATATCTGGCTTAAAGAAAGAGAGGTCTTATGCCAGGACTCCTTTTTC 660
Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 661 ATGAACTCATCAGCAGATGAGAGCTTCTACTGTGACTTTGCTGCTGCTGATGTTCCAA 720
Qy 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
Db 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGTGAGGAGATCATTGTTGATGCTGTCMAA 780
Qy 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 781 ATTGAGCAGAGTATTAACAGAGGCTTGGCCAGTGGGCTCATTTGGAATGAATTCATT 840
Qy 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 841 TTGATGAACAGTACATCAGATTGAGTTGAGCTGACAGATTAATCTTGGAACTTGGATTTCA 900
Qy 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 901 AAGGTTTTTCAGCAGAAAATCCTTTTGAATTTATGGAACACATTTCTTTAGAGGAAA 960
Qy 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 961 ACAAAATTTCTTGGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGCTTATGGCAGAAACC 1020
Qy 341 ThrAsnAsnValPheThrLeuAspAlaAspPhe 351
Db 1021 ACAGATAACGTCTTCACCTTGGATGAGATTTT 1053

RESULT 3

US-10-698-228-4
; Sequence 4, Application US/10698228
; Publication No. US2004007253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; PRIOR FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391

; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 4

; LENGTH: 1081
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:

US-10-698-228-4

Alignment Scores:
Pred. No.: 1,62e-216 Length: 1081
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 19 Gaps: 0

US-10-698-228-1 (1-351) x US-10-698-228-4 (1-1081)

Qy 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
Db 20 ATGGCGCACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCACTTCA 79
Qy 21 AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerSerArg 40
Db 80 GACACCAACGAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 139
Qy 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
Db 140 CGGTTTGTCACTTTTCCAATCCAGTACCTCATATTTGGAATAATGATAAACAGGCACAG 199
Qy 61 AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
Db 200 GCITTCCTTCTGGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCAGTGAACAAG 259
Qy 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Db 260 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTACATCTTAGCCTTTTTCAGCCAGCTGAT 319
Qy 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 320 GSAATGTTAAATGAAATTTGGTGGAGCGCTTTTAGTCAGAGAGTGCAGGTTCCAGAGGCT 379
Qy 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 380 CGCTGTTCTATGGCTTTCAAAATTTCTATCAGAGAAATGTTCACTCAGAGATGTACAGTTG 439
Qy 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 440 CTGATAGACATTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAATGCAATGAA 499
Qy 161 ThrMetProTyrValLysLysLysAlaAspTyrPalaLeuArgTyrIleAlaAspArgLys 180
Db 500 ACCATGCCCTTATGTTAAGAAAAAAGCAGATTGGCCCTTGCATGGATAGCAGATAGAAA 559
Qy 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
Db 560 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAGAGAGTTTCTTCTCAGGA 619
Qy 201 SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 620 TCTTTTGGCTGCTATATCTTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCCTTTTTC 679
Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 680 AATGAACTCATCAGCAGAGATGAAGGACTTTCACCTGTGACTTTGCTGCTGCTGATGTTCAA 739
Qy 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
Db 740 TACTTAGTAAATAAGCCTTCAGAAAGAGGCTCAGGAGATCATTTGTTGATGCTGCTCAA 799


```
QY 261 IleGluGlnPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysile 280
|||
Db 800 ATTGAGCAGAGTTTAAACAGAGCTTGCAGTTGGCCTCATTTGAATGAATTCATT 859
|||
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
|||
Db 860 TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 919
|||
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
|||
Db 920 AAGGTTTTTTCAGCAGAAATCTTTTGAATTTATGAGAAACATTTCTTTAGAGGAAAA 979
|||
QY 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
|||
Db 980 ACAAAATTTCTTGAGAAACAGTTTTCAGAGTATCAGCGTTTGCAGTTATGGCAGAAACC 1039
|||
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
|||
Db 1040 ACAGATAACGTCTTCACTTGGATGCAGATTTT 1072
|||
RESULT 4
US-10-172-118-71
; Sequence 71, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 71
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AB036063
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-71
Alignment Scores:
Pred. No.: 1,77e-215 Length: 4955
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0
US-10-698-228-1 (1-351) x US-10-172-118-71 (1-4955)
QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
|||
Db 245 ATGGGCGACCCGGAAGGCGGAGCGCGGCTGGATCAGATCAGAGATCATCTTCA 304
|||
QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
|||
Db 305 GACACCAACGAAGTGAATTAAGTCAATGAAGACCACTCTTAAGAAAGAGTTCTCGC 364
|||
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
|||
Db 365 CGGTTTGTTCATCTTCCATCCAGTACCTGATATTTGGAAATGTATAAACAGGCACAG 424
|||
QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
|||
```

```
Db 425 GCTTCCTCTTGGACAGCAGAGAGGTGCAGCTTATCAAGAGATCTCCCTCACTGGAACAAG 484
|||
QY 81 LeuLysAlaAspGlnLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
|||
Db 485 CTTAAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTAGCCTTTTTCAGCAGCAGTAT 544
|||
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
|||
Db 545 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGTTTCCAGAGCT 604
|||
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
|||
Db 605 CGCTGTTTCTATGGCTTCAATTCATCAGAAATCTTCACTCAGAGATGTACAGTTTG 664
|||
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
|||
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTCGA 724
|||
QY 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
|||
Db 725 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCCCTTGGCATGGATAGCAGATAGAAA 784
|||
QY 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
|||
Db 785 TCTACTTTTGGGGAAGAGTGTGGCTTGTCTGTAGAGAGAGTTTCTTCTCAGGA 844
|||
QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
|||
Db 845 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACATTTTTC 904
|||
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
|||
Db 905 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGTCTGCTGATGTCCA 964
|||
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleLeuValAspAlaValLys 260
|||
Db 965 TACTTAGTAATTAAGCCTTCAGAAAGAGGTGAGGAGATCATTTGTGCTGTCTCA 1024
|||
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysile 280
|||
Db 1025 ATTGAGCAGGAGTTTTTAAACAGAGCCTTCCAGTTGGCCTCATTTGGAATGAATTCAT 1084
|||
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
|||
Db 1085 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTACTTTGGAACTTGGATTTCTCA 1144
|||
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
|||
Db 1145 AAGTTTTTTCAGCAGAAATCTTTTGAATTTATGGAACATTTCTTTAGAGGAAAA 1204
|||
QY 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
|||
Db 1205 ACAATTTTCTTTCAGAAACAGTTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1264
|||
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
|||
Db 1265 ACAGATAACGTCTTCACTTGGATGCAGATTTT 1297
|||
RESULT 5
US-10-342-887-71
; Sequence 71, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
```



```
QY 61 AlaSerPheThrAlaGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
Db 425 GCTTCTCTCTGACAGCAGAGAGGTGCGACTTATCAAGAGTCTCCCTCCTCGAACAAG 484
QY 81 LeuLysAlaAspGluLysTyrPheLeSerHisLeLeuAlaPhePheAlaAlaSerAsp 100
Db 485 CTTAAAGCAGATGAGAAGTACTTCTCTCATCTCTACATCTTAGGCTTTTTCAGCCAGTGAT 544
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 545 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGTTCAGAGGCT 604
QY 121 ArgCysPheThrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 605 CGCTGTTTCTATGCTTTCCTCAATTCATCGAAGTGTCTCAGAGATGTACAGTTTG 664
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 724
QY 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Db 725 ACCATGCCCTATGTTAAAGAAAAAGCAGATTGGGCTTGGCATGGATAGCAGATAGAAAA 784
QY 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
Db 785 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGTAGAGAGGTGTTCTTCTCAGGA 844
QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 845 TCTTTTGTGCTATATTCTGGCTAAAGAAAGAGAGGTCTTATGCCAGGACTCACTTTTTC 904
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 905 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 964
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleLeuValAspAlaValLys 260
Db 965 TACTTAGTAAATTAAGCTTTTCAAGAAAGAGGTGAGGAGATCATTTGTTGATGCTGTCAA 1024
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 1025 ATTGACGAGAGTTTTACAGAGCTTGGCAGTGGCTCATTTGAATGAATTCATTT 1084
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 1085 TTGATGAACAGTACATTGAGTTGTAGCTGACAGATTACTTGTGAACTTGTGATCTCA 1144
QY 301 LysValPheGluAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 1145 AAGGTTTTTTCAGCAGAAAATCTTTTGAATTTTATGGAATAACATTTCTTTAGAAAGAAA 1204
QY 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 1205 ACAATTTCTTTCAGAAAACAGTTTCAGAGTATCAGCGTTTTCAGATTATGCGAAGAAC 1264
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1265 ACAGATAACGTCTTCACTTGGATGAGATTTT 1297
```

RESULT 7

```
US-09-925-301-505
; Sequence 505, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
```

```
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 505
; LENGTH: 1989
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1917)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-505
```

```
Alignment Scores:
Pred. No.: 4,56e-171 Length: 1989
Score: 1460.50 Matches: 277
Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
Gaps: 9
Db:
```

US-10-698-228-1 (1-351) x US-09-925-301-505 (1-1989)

```
QY 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
Db 185 CTTGACGGGACCGCGCTGCTGGCCAGCAAGACCGCAGGAGGATCTTCCAGGAGCCAC 244
QY 23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
Db 245 GGAGCGCAAACTAAAGCAGCTGCCCGCGCTGGAGGATGAGCGCTGCTCAGAGAAA 304
QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysG 58
Db 305 CCCCCCGCTTTGTCTATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 364
QY 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
Db 365 GGACAGAGCTTCTTTTGGACCGCCGAGGAGGTGGACCTTCCAGAGGACATTCAGCATG 424
QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisLeuAlaPhePheAlaAl 98
Db 425 GGAATCCCTGAAACCCCGAGGAGAGATATTTATATCCATGCTTCTGGCTTCTTTGCAGC 484
QY 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
Db 485 AAGCGATGGCATAGTAATGAAACCTTGGTGGAGCGATTAGCCCAAGAAAGTTTCAGATTAC 544
QY 118 oGluAlaAtgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
Db 545 AGAAGCCCGCTGTTCTATGGCTTCCAAATTCATGCAATGCAATGCAATGCAATGTA 604
QY 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
Db 605 TAGTCTCTTTATTGACACTTACATAAAAGATCCCAAGAAAGGAAATTTCTTCAATGC 664
QY 158 alIeGluThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
Db 665 CATTGAAACGATGCTTGTGTCAGAAAGAGGAGACTGGGCGCTTGGCGCTGGATTGGGGA 724
QY 178 pArgLysSerThrPheGlyGluArgValAlaAlaPheAlaAlaValGluGlyValPhePh 198
Db 725 CAAAGAGGCTACCTATGTTGAAACGTGTTGATAGCTTTGCTCAGTGGAGGCAATTTCTT 784
QY 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
Db 785 TTCGGTCTCTTTGGTCGATATTTCTGCTCAAGAACGAGACTGATGCTCGGCTCAC 844
QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
Db 845 ATTTTCTAATGAACCTTATTAGCAGAGATGAGGGTTTACACTGTGATTTTGTGCTGAT 904
QY 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
```


Db 302 CCTGAGCGGACCCCGCTCTGCGCCAGCAAGACCGCGAGGAGATCTTCCAGGAGCCAC 361
QY 23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
Db 362 GGAGCGGAAACTAAAGAGCTGCCCGCGCGTGGAGGATGAGCGCTGCTGAGAGAAA 421
QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysG1 58
Db 422 CCCCCCGCGCTTGTCTATCTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481
QY 58 nAlaGlnAlaSerPheTrpThrAlaGluValAspLeuSerLysAspLeuProHisTr 78
Db 482 GGAGAGGCTTCTTTTGACCGCGAGAGGTGTACCTCTCCAGGACATTCAGCACTG 541
QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAl 98
Db 542 GGAATCCCTGAAACCCCGAGGAGATATTTATATCCATGTTCTGGCTTCTTTGAGC 601
QY 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
Db 602 AAGCGATGCGATAGTAATGAAACTTGTGTGAGCGATTTAGCCAGAGTTCAGATTAC 661
QY 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
Db 662 AGAAGCGCGCTTCTTATGGCTTCCAAATGCCAGAAACATACATTCCTGAAATGTA 721
QY 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
Db 722 TAGTCTTCTTATGACACTTACATAAAGATCCCAAGAAAGGGAATTTCTCTTCATGC 781
QY 158 aIleGluThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
Db 782 CATTGAACGATGCTTGTGTCAAGAGAGGAGCAGACTGGCGCTTCGCTGGATTCGGGA 841
QY 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePh 198
Db 842 CAAAGAGGCTACTATGGTGAACGTGTGTAGCCTTGTGCTGAGTGGAGGCCATTTTCT 901
QY 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
Db 902 TTCCGGGTCTTTTGGTTCGATATTCCTGGCTCAAGAAACGAGGACTGATCGCTGCCCTAC 961
QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
Db 962 ATTTCTAATGACTTATTAGCAGATGAGGTTTACACTGTGATTTTGTCTGCTGAT 1021
QY 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
Db 1022 GTTCAACACCTGTTACACAAACCATCGAGGAGAGATGAGAGAAATTAATTATCAATGC 1081
QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
Db 1082 TGTTCCGATAGAACAGGAGTTCTCTCACTGAGCGCTTGCCTGTGAAGCTCATTTGGATGAA 1141
QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGl 298
Db 1142 TTGCACTCTAATGAAGCAATATCATTTAGTTTGTGGCAGACAGACTTATCTGGAACCTGG 1201
QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGl 318
Db 1202 TTTTAGCAAGGTTTTCAGAGTAGAGAACCCATTTGACITTTATGGAGAAATTTTCACTGGA 1261
QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
Db 1262 AGGAAAGACTAACTTCTTTGAGAGAGAGTAGGCGAGTATCAGAGGATGGGAGTGATGTC 1321
QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1322 AAGTCCACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1361

RESULT 11

US-09-954-456-1169

; Sequence 1169, Application US/09954456

; Patent NO. US20020115057A1

; GENERAL INFORMATION:

; APPLICANT: Young, Paul

; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Can

; TITLE OF INVENTION: Sets

; FILE REFERENCE: 689290-76

; CURRENT APPLICATION NUMBER: US/09/954,456

; CURRENT FILING DATE: 2001-09-18

; PRIOR APPLICATION NUMBER: US/60/233,617

; PRIOR FILING DATE: 2000-09-18

; PRIOR APPLICATION NUMBER: US/60/234,052

; PRIOR FILING DATE: 2000-09-20

; PRIOR APPLICATION NUMBER: US/60/234,923

; PRIOR FILING DATE: 2000-09-25

; PRIOR APPLICATION NUMBER: US/60/235,134

; PRIOR FILING DATE: 2000-09-25

; PRIOR APPLICATION NUMBER: US/60/235,637

; PRIOR FILING DATE: 2000-09-26

; PRIOR APPLICATION NUMBER: US/60/235,638

; PRIOR FILING DATE: 2000-09-26

; PRIOR APPLICATION NUMBER: US/60/235,711

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: US/60/235,720

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: US/60/235,840

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: US/60/235,863

; PRIOR FILING DATE: 2000-09-27

; NUMBER OF SEQ ID NOS: 2276

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 1169

; LENGTH: 2500

; TYPE: DNA

; ORGANISM: Homo sapiens

; US-09-954-456-1169

Alignment Scores:
Pred. No.: 6,53e-171 Length: 2500
Score: 1460.50 Matches: 277
Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
DB: 9 Gaps: 1

US-10-698-228-1 (1-351) x US-09-954-456-1169 (1-2500)

QY 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
Db 302 CCTGAGCGGACCCCGCTCTGCGCCAGCAAGACCGCGAGGAGATCTTCCAGGAGCCAC 361
QY 23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
Db 362 GGAGCGGAAACTAAAGAGCTGCCCGCGCGTGGAGGATGAGCGCTGCTGAGAGAAA 421
QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysG1 58
Db 422 CCCCCCGCGCTTGTCTATCTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481
QY 58 nAlaGlnAlaSerPheTrpThrAlaGluValAspLeuSerLysAspLeuProHisTr 78
Db 482 GGAGAGGCTTCTTTTGACCGCGAGGAGGTGTACCTCTCCAGGACATTCAGCACTG 541
QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAl 98
Db 542 GGAATCCCTGAAACCCCGAGGAGATATTTATATCCATGTTCTGGCTTCTTTGAGC 601
QY 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
Db 602 AAGCGATGCGATAGTAATGAAACTTGTGGAGCGGATTTAGCCAGAGATTCAGATTAC 661
QY 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
Db 662 AGAAGCGCGCTTCTTATGGCTTCCAAATGCCAGAAACATACATTCCTGAAATGTA 721

QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
 Db 1082 TGTTCGATAGAACAGAGTTCTCTACTGAGGCTTGGCTGTGAAGCTCATTTGGATGA 1141
 QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuG 298
 Db 1142 TTGCATCTTAATGAACAATACATTGAGTTGTGGCAGACAGACTTATGCTGGAACCTGGG 1201
 QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG 318
 Db 1202 TTTTACAGAGTTTTCAGAGTAGAGAACCCATTGACTTTATGGAGAAATATTCACATGGA 1261
 QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
 Db 1262 AGGAAGAGCTAACTCTTTTGGAGAAGAGAGTAGGCGAGATATCAGAGGATGGAGTGATGTC 1321
 QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
 Db 1322 AAGTCCACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1361

RESULT 13

US-10-641-643-1370

; Sequence 1370, Application US/10641643

; Publication No. US20040077003A1

; GENERAL INFORMATION:

; APPLICANT: Cocks, Benjamin G.

; Susan G. Stuart

; Jeffrey J. Seilhamer

; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL

; GENE EXPRESSION

; NUMBER OF SEQUENCES: 1508

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

; STREET: 3174 PORTER DRIVE

; CITY: PALO ALTO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

; CURRENT APPLICATION DATA: US/10/641,643

; FILING DATE: 14-Aug-2003

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: <Unknown>

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Zeller, Karen J.

; REGISTRATION NUMBER: 37,071

; REFERENCE/DOCKET NUMBER: PA-0001 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (650) 855-0555

; TELEFAX: (650) 845-4166

; INFORMATION FOR SEQ ID NO: 1370:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2500 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; IMMEDIATE SOURCE:

; LIBRARY: GENBANK

; CLONE: g36154

; SEQUENCE DESCRIPTION: SEQ ID NO: 1370 :

US-10-641-643-1370

Alignment Scores:

Pred. No.: 6,53e-171 Length: 2500

Score: 1460.50 Matches: 277

Percent Similarity: 88.42% Conservative: 36

Best Local Similarity: 78.25% Mismatches: 35
 Query Match: 80.20% Indels: 6
 DB: 19 Gaps: 1

US-10-698-228-1 (1-351) x US-10-641-643-1370 (1-2500)

QY 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
 Db 302 CCTGAGCGGACCCCGCTCTGGCCAGCAGCAGCGGAGGAGATCTTCCAGGAGCCAC 361
 QY 23 nGluSerGluIleLysSer-----AsnGluGluProLeuLeuArgLysSe 38
 Db 362 GGAGCCGAAAACTAAAGCAGCTGCCCGCGCGTGGAGGATGAGCGCTGCTCAGAGAAA 421
 QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysG 58
 Db 422 CCCCCGCGCTTGTCTATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAAGAA 481
 QY 58 nAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHistr 78
 Db 482 GGCAGAGGCTTCTTTTGGACCGCGGAGGAGTTGACCTCTCCAGGACATTCACACTG 541
 QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAl 98
 Db 542 GGAATCCCTGAAACCCCGAGGAGAGATATTTATATCCATGTTCTGGCTTCTTTTCAGC 601
 QY 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
 Db 602 AAGCGATGGCATAGTAAATGAAACCTTGGTGGAGCGATTTAGCCAAAGAAAGTTTCAGATT 661
 QY 118 oGluAlaAatGcysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
 Db 662 AGAAGCCCGCTGTTTCTATGCTTCCAAATTTGCCATGAAACATACATTTCTGAATGTA 721
 QY 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
 Db 722 TAGTCTTCTTATTGACACTTACATAAAGATCCCAAGAAAGGGAATTTCTTCTCAATGC 781
 QY 158 alIeGluThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
 Db 782 CATTAAGAACGATCGCTTGTGTCAAGAAAGAGGAGAGCTGGGCTTGGCTGGATGGGA 841
 QY 178 pArgLysSerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePh 198
 Db 842 CAAAGAGGCTACCTATGTTGAACGTGTGTAGCTTGTGCTGAGTGAAGGACATTTCTT 901
 QY 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
 Db 902 TTCGGTCTTTTGGCTGATATTTGGCTCAAGAAACGAGGACTGATGCTGGCTCAC 961
 QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
 Db 962 ATTTCTAATGAACCTTATTAGCAGAGATGAGGGTTTACACTGTGATTTGCTTGCCTCAT 1021
 QY 238 tPheGlnTyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAl 258
 Db 1022 GTTCAACACCTGGTACACAAACCATCGGAGGAGAGATGAAGAAAATAATATCAATGC 1081
 QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
 Db 1082 TGTTCGATAGAACAGAGGTTCTCTACTGAGGCTTGCCTGTGAGCTCATTTGGATGAA 1141
 QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuG 298
 Db 1142 TTGCATCTTAATGAAGCAATACATTGAGTTTGTGGCAGACAGACTTATGCTGGAACCTGG 1201
 QY 298 yPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuG 318
 Db 1202 TTTTAGCAAGGTTTTCAGAGTAGAGAACCCATTGACTTATGGAGAAATATTTTCACATG 1261
 QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
 Db 1262 AGGAAGAGCTAACTCTTTTGGAGAAGAGTAGGCGAGATATCAGAGGATGGAGTGATGTC 1321


```
QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1322 AAGTCCACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1361

RESULT 14
US-10-733-878-458
; Sequence 458, Application US/10733878
; Publication No. US20040224408A1
; GENERAL INFORMATION:
; APPLICANT: Jean-Philippe Girard
; APPLICANT: Francois Amalric
; APPLICANT: Myriam Roussigne
; APPLICANT: Thomas Clouaire
; TITLE OF INVENTION: THAP PROTEINS AS NUCLEAR RECEPTORS FOR
; TITLE OF INVENTION: CHEMOKINES AND ROLES IN TRANSCRIPTIONAL REGULATION, CELL
; TITLE OF INVENTION: PROLIFERATION AND CELL DIFFERENTIATION
; FILE REFERENCE: BIOBANK.012A
; CURRENT APPLICATION NUMBER: US/10/733,878
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 60/432699
; PRIOR FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/485027
; PRIOR FILING DATE: 2003-07-03
; NUMBER OF SEQ ID NOS: 535
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 458
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-733-878-458

Alignment Scores:
Pred. No.: 6,53e-171 Length: 2500
Score: 1460.50 Matches: 277
Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
DB: 21 Gaps: 1

US-10-698-228-1 (1-351) x US-10-733-878-458 (1-2500)

QY 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
Db 302 CTGAGCGGAGCCCGCTCTGCCAGCAAGACCGGAGAGGATCTTCAGGAGGCCAC 361

QY 23 nGluSerGluLeuLeuSer-----AsnGluGluProLeuLeuArgLysSe 38
Db 362 GGAGCGGAAAACTAAAGCAGTGCCTCCCGCGGTGGAGGATGAGCCGCTGCTGAGAGAAAA 421

QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTTrpLysMetTyrLysG 58
Db 422 CCCCCCGCGCTTGTCTATCTTCCCATCGAGTACCATGATATCTGGCAGATGTATAGAA 481

QY 58 nAlaGlnAlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisStr 78
Db 482 GCGAGAGGCTTCTTTTGGACCGCCGAGGAGGTGTACCTCTCAAGACATTCAGCAGCTG 541

QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAl 98
Db 542 GGAATCCCTGAAACCGGAGAGAGATATTTATATCCATGTTCTGGCTTTCTTTCAGC 601

QY 98 aSerAspGlyLeValAsnGlnAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
Db 602 AAGCGATGCGATAGTAATAATGAAACCTTGGTGGAGGATTTAGCCAAAGAGTTCCAGATTAC 661

QY 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuLeuGluAsnValHisSerGluMetTy 138
Db 662 AGAAGCCCGCTGTTCTATGGCTTCCAAATTTGCCATGGGAAAAATACATCTTCAAAATGTA 721

QY 138 rSerLeuLeuLeuAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
Db 722 TAGTCTTCTTATGACACTTACATATAAAGATCCCAAGAAAGAGGAATTTCTTCTCATGTC 781
```

; NUMBER OF SEQ ID NOS: 8447
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3751
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-843-641A-3751

Alignment Scores:
Pred. No.: 6,538-171 Length: 2500
Score: 1460.50 Matches: 277
Percent Similarity: 88.42% Conservative: 36
Best Local Similarity: 78.25% Mismatches: 35
Query Match: 80.20% Indels: 6
DB: 22 Gaps: 1

US-10-698-228-1 (1-351) x US-10-843-641A-3751 (1-2500)

```
QY 4 ProGluArgProGluAlaAlaGlyLeuAspGlnAspGlu-ArgSerSerSerAspThrAs 23
DB 302 CCTGAGCGGGACCCCGCTCTGGCCAGCAGACCGGAGGAGATCTCCAGGAGCCAC 361
QY 23 nGluSerGluLeuLeuSer-----AsnGluGluProLeuLeuArgLysSe 38
DB 362 GGAGCCGAAACTAAAGCAGCTGCCCGCGCTGGAGGATGAGCGCTGCTGAGAGAAA 421
QY 38 rSerArgArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGl 58
DB 422 CCCCAGCGCGCTTGTTCATCTCCCATCGAGTACCATGATATCTGGCAGATGATAAGAA 481
QY 58 nAlaGlnAlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTr 78
DB 482 GCGAGAGGCTTCCTTTTGACCGCGGAGGAGTTGACCTCTCCAAGGACATTCAGCAGTG 541
QY 78 pAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAl 98
DB 542 GGAATCCCTGAAACCGGAGGAGATATTTATATCCCATGTTCTGGCTTTCTTTCGACG 601
QY 98 aSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValPr 118
DB 602 AAGCGATGGCATAGTAATGAAACTTGGTGAGCGCATTTAGCCAAAGAAAGTTTCAGATTAC 661
QY 118 oGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTy 138
DB 662 AGAAGCCCGCTGTTCTATGGCTTCCAAATGGCCATGGAAACATACATCTCGAAATGTA 721
QY 138 rSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAl 158
DB 722 TAGTCTTCTTATTGACACTTACATNAAGATCCCAAGAGAGGGAATTTCTTTCATGTC 781
QY 158 aIleGluThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAs 178
DB 782 CATTGAACAGCATGCTTGTGTCAAGAGAAGCAGACTGGGCTTGGCTGGATTGGGA 841
QY 178 pArgLysSerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePh 198
DB 842 CAAGAGGCTACCTATGTGAACGTGTTGTAGCCCTTGTGTCAGTGAAGGCAATTTCTT 901
QY 198 eSerGlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuTh 218
DB 902 TTCGGTTCCTTTGGCTCGATATCTGGCTCAAGAAACGAGGACTGATGGCTGCCTCAC 961
QY 218 rPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMe 238
DB 962 ATTTCTTAATGAACCTTATTAGCAGAGATGAGGGTTTACATGTGATTTTGTCTGCTGAT 1021
QY 238 tPheGlnTyrLeuValAsnLysProSerGluArgValArgGluIleIleValAspAl 258
DB 1022 GTTCAAAACACCTGGTACCAAAACCATCGGAGGAGAGTAAGAGAAATAATTTATCAATGC 1081
QY 258 aValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAs 278
DB 1082 TGTTCGGATAGACAGGAGTTCCTCCTGAGGCCCTTGTGAGGCTCATTTGGGATGAA 1141
```

```
QY 278 nCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuG1 298
DB 1142 TTGCACCTCTAATGAGCAATACATTGAGTTTGTGGCAGACAGACTTATGCTGGAACTGGG 1201
QY 298 yPheSerLysValPheGlnAlaGluAsnPropheAspPheMetGluAsnIleSerLeuG1 318
DB 1202 TTTTAGCAAGGTTTTCAGAGTAGAGAACCCATTTGACTTTTATGGAGAATAATTTCACTGGA 1261
QY 318 uGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAl 338
DB 1262 AGGAAAGACTAACTCTCTTTGAGAAGAGAGTAGGCGAGTATCAGAGGATGGGAGTGATGTC 1321
QY 338 aGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB 1322 AAGTCCAACACAGAGAAATCTTTTACCTTGGATGCTGACTTC 1361
```

Search completed: October 30, 2005, 10:18:11
Job time : 679 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2005, 21:42:15 ; Search time 3258 Seconds
(without alignments)
15660.940 Million cell updates/sec

Title: US-10-698-228-2

Perfect score: 1053

Sequence: 1 atggcgaccgcgaaagcc.....tcaccttgatgcagatttt 1053

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1053	100.0	1053	6	AR454868 Sequence
2	1053	100.0	1053	6	BD064764 Novel pro
3	1053	100.0	1053	6	BD093077 Novel pro
4	1053	100.0	1056	6	CQ714252 Sequence
5	1053	100.0	1056	9	AB166671 Homo sapi
6	1053	100.0	4955	6	AR454869 Sequence
7	1053	100.0	4955	6	BD064765 Novel pro
8	1053	100.0	4955	6	BD093078 Novel pro
9	1053	100.0	4955	9	AB036063 Homo sapi
10	1051.4	99.8	1053	6	AR454877 Sequence
11	1051.4	99.8	1053	6	BD064773 Novel pro
12	1051.4	99.8	1053	6	BD093086 Novel pro
13	1051.4	99.8	1081	6	AR454870 Sequence
14	1051.4	99.8	1081	6	BD064766 Novel pro
15	1051.4	99.8	1081	6	BD093079 Novel pro
16	1049.8	99.7	1601	6	BD156916 Primer fo
17	1049.8	99.7	1601	6	AX877905 Sequence
18	1049.8	99.7	1601	9	AK001965 Homo sapi
19	895.2	85.0	4532	10	BC058103 Mus muscu

20	855	81.2	4571	9	BC042468	BC042468 Homo sapi
21	853.4	81.0	900	9	AB163437	AB163437 Homo sapi
22	837.6	79.5	171737	10	AC122379	AC122379 Mus muscu
23	579	55.0	2500	6	AR300103	AR300103 Sequence
24	579	55.0	2500	6	AR438503	AR438503 Sequence
25	579	55.0	2500	6	AR562011	AR562011 Sequence
26	578.4	54.9	1649	9	BC001886	BC001886 Homo sapi
27	578.4	54.9	1653	9	BC030154	BC030154 Homo sapi
28	578.4	54.9	1794	9	SA0301	SA0301 Homo sapien
29	578.4	54.9	2500	6	AR380825	AR380825 Sequence
30	578.4	54.9	2500	6	AX333242	AX333242 Sequence
31	578.4	54.9	2500	6	AX333687	AX333687 Sequence
32	578.4	54.9	2500	6	AX334345	AX334345 Sequence
33	578.4	54.9	2500	9	HSRR25S	X59618 H.sapiens R
34	578.4	54.9	3039	9	AK092671	AK092671 Homo sapi
35	578.2	54.9	1562	5	BC061353	BC061353 Xenopus t
36	574.6	54.6	1538	5	CR762018	CR762018 Xenopus t
37	574.6	54.6	1611	5	CR848103	CR848103 Xenopus t
38	574.6	54.6	1647	5	BC080161	BC080161 Xenopus t
39	573.6	54.5	217774	2	AC120603	AC120603 Rattus no
40	573.2	54.4	3203	9	AK123010	AK123010 Homo sapi
41	570.4	54.2	2111	10	MUSRNM2A	M14223 Mouse ribon
42	568.8	54.0	1170	6	E14825	E14825 cDNA encodi
43	566.2	53.8	1570	5	BC072071	BC072071 Xenopus l
44	562.6	53.4	2057	5	BC044355	BC044355 Danio rer
45	561.8	53.4	1406	5	CR353642	CR353642 Gallus ga

ALIGNMENTS

RESULT 1	AR454868	AR454868	Sequence 2 from patent US 6682917.	1053 bp	DNA	linear	PAT 20-FEB-2004
LOCUS	AR454868	Sequence 2 from patent US 6682917.					
DEFINITION	AR454868	Sequence 2 from patent US 6682917.					
ACCESSION	AR454868	Sequence 2 from patent US 6682917.					
VERSION	AR454868.1	GI:42688823					
KEYWORDS	Unknown.						
SOURCE	Unknown.						
ORGANISM	Unknown.						
REFERENCE	1 (bases 1 to 1053)						
AUTHORS	Nakamura, Y., Arakawa, H. and Tanaka, H.						
TITLE	Protein having a ribonucleotide Reductase activity and a DNA thereof						
JOURNAL	Patent: US 6682917-A 2 27-JAN-2004;						
FEATURES	Location/Qualifiers						
source	1..1053						
ORIGIN	/organism="unknown"						
	/mol_type="genomic DNA"						

Query Match	100.0%;	Score 1053;	DB 6;	Length 1053;
Best Local Similarity	100.0%;	Pred. No. 1.2e-255;		
Matches 1053;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	ATGGCGACCCCGAAAGCGCGGCTGGATCAGATGAGATGAGATCATCTTCA	60	
Db	1	ATGGCGACCCCGAAAGCGCGGCTGGATCAGATGAGATGAGATCATCTTCA	60	
Qy	61	GACACCAACGAAGTGAATTAAGTCAATTAAGAGCCACTCTTAAGAAAGTTCTCGC	120	
Db	61	GACACCAACGAAGTGAATTAAGTCAATTAAGAGCCACTCTTAAGAAAGTTCTCGC	120	
Qy	121	CGTTTGTTCATCTTCCCAATCCAGTACCTCGATATTTGGAAATGTATAAAGCAGCAG	180	
Db	121	CGTTTGTTCATCTTCCCAATCCAGTACCTCGATATTTGGAAATGTATAAAGCAGCAG	180	
Qy	181	GCTTCCTTCTGGACAGCAGAGAGGTCGATTAATCAAGAGTCTCCCTCACTGGAAACAAG	240	
Db	181	GCTTCCTTCTGGACAGCAGAGAGGTCGATTAATCAAGAGTCTCCCTCACTGGAAACAAG	240	
Qy	241	CTTAAGACAGATGAGAGTACTTCTCTCACATCTTAGCCTTTTGGAGCCAGTGTAT	300	

Db 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCACTCTTAGCCCTTTTTCGAGCCAGTGAT 300
QY 301 GGAATTTGTAATGAAAATTTGGTGGAGCCCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
Db 301 GGAATTTGTAATGAAAATTTGGTGGAGCCCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
QY 361 CGCTGTTCTATGCGCTTCAAAATTCATCGAGAAATGTTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTCTATGCGCTTCAAAATTCATCGAGAAATGTTTCACTCAGAGATGTACAGTTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTATTAATGCAATTTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTATTAATGCAATTTGAA 480
QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATGCGGCTTCGATGGATAGCAGATAGAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAAGCAGATGCGGCTTCGATGGATAGCAGATAGAAA 540
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTTCTGCTGTAAGAGAGGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTTCTGCTGTAAGAGAGGTTTCTTCTCAGGA 600
QY 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGGTTTATGCCAGGACTCACTTTTTC 660
Db 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGGTTTATGCCAGGACTCACTTTTTC 660
QY 661 AATGAACCTCATCAGCAGATGAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 661 AATGAACCTCATCAGCAGATGAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
QY 721 TACTTAGTAAATAAGCTTTCAGAAAGAGGTCAGGAGATCACTTGTGATGCTGCAAA 780
Db 721 TACTTAGTAAATAAGCTTTCAGAAAGAGGTCAGGAGATCACTTGTGATGCTGCAAA 780
QY 781 ATTGAGCAGAGTTTTAAACAGAAAGCTTGCAGAGTGGCTCATTTGGAATGAATTCATT 840
Db 781 ATTGAGCAGAGTTTTAAACAGAAAGCTTGCAGAGTGGCTCATTTGGAATGAATTCATT 840
QY 841 TTGATGAAACAGTACATGAGTTTGTAGCTGACAGATTACTTGTGAACTTGGAACTCTCA 900
Db 841 TTGATGAAACAGTACATGAGTTTGTAGCTGACAGATTACTTGTGAACTTGGAACTCTCA 900
QY 901 AGGTTTTTCAGCAGAAATCCTTTTGAATTTATGGAACATTTCTTTAGAGGAAA 960
Db 901 AGGTTTTTCAGCAGAAATCCTTTTGAATTTATGGAACATTTCTTTAGAGGAAA 960
QY 961 ACAAAATTTCTTTGAGAAAAGGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020
Db 961 ACAAAATTTCTTTGAGAAAAGGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020
QY 1021 ACAGATAACGTTCTTCACTTGGATGAGATTTT 1053
Db 1021 ACAGATAACGTTCTTCACTTGGATGAGATTTT 1053

RESULT 2
BD064764
LOCUS
DEFINITION Novel protein and DNA thereof.
ACCESSION BD064764
VERSION BD064764.1 GI:22610367
KEYWORDS JP 2001269184-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1053)
REFERENCE
AUTHORS Nakamura, Y., Arakawa, H. and Tanaka, H.
TITLE Novel protein and DNA thereof
JOURNAL Patent: JP 2001269184-A 1 02-OCT-2001;
YUSUKE NAKAMURA, TAKEDA CHEMICAL INDUSTRIES LTD
COMMENT OS Homo sapiens (human)

PN JP 2001269184-A/1
PD 02-OCT-2001
PI 27-JUN-2000 JP 2000192401
PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
C12N15/09, A61K31/711, A61K38/43, A61K45/00, A61K48/00, PC
A61P35/00,
PC A61P43/00, C07K16/32, C07K16/40, C12N1/15, C12N1/19, C12N1/21 PC
C12N5/10, C12N9/02,
PC C12P21/02, C12P21/08, G01N33/15, G01N33/50, G01N33/53// (C12N15/09,
C12R1:91),
PC (C12N15/09, C12R1:19), (C12N9/02, C12R1:91), (C12P21/02, C12R1:91),
PC (C12P21/08, C12R1:91), (C12N15/00, A61K37/02, A61K37/48, C12N5/00,
PC (C12N15/00, C12R1:91), (C12N15/00, C12R1:19)
CC Novel protein and DNA thereof
FH Key Location/Qualifiers
FT source 1..1053 /organism="Homo sapiens (human)".
FT source Location/Qualifiers
1..1053
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
FEATURES
source
ORIGIN
Query Match 100.0%; Score 1053; DB 6; Length 1053;
Best Local Similarity 100.0%; Pred. No. 1.2e-255;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCGACCGGAAAGCGCGGAGCGGCTGGATCAGATGAGAGATCATCTTCA 60
Db 1 ATGGCGACCGGAAAGCGCGGAGCGGCTGGATCAGATGAGAGATCATCTTCA 60
QY 61 GACACCAACAAAGTGAATTAAGTCAATCAAGAGCCACTCTTAAGAAAGGTTCTGC 120
Db 61 GACACCAACAAAGTGAATTAAGTCAATCAAGAGCCACTCTTAAGAAAGGTTCTGC 120
QY 121 CGTTTGTCTATCTTCCAAATCCAGTACCTGATATTTGGAAAATGTATAACAGGCACAG 180
Db 121 CGTTTGTCTATCTTCCAAATCCAGTACCTGATATTTGGAAAATGTATAACAGGCACAG 180
QY 181 GCTTCTCTTGGCAGCAGAGAGGTCGACTTATCAAGAGTCTCCCTCACTGGAAACAG 240
Db 181 GCTTCTCTTGGCAGCAGAGAGGTCGACTTATCAAGAGTCTCCCTCACTGGAAACAG 240
QY 241 CTTAAGCAGATGAGAGTACTTCTCATCTCAGATCTTAGCCTTTTTCGAGCCAGTGTAT 300
Db 241 CTTAAGCAGATGAGAGTACTTCTCATCTCAGATCTTAGCCTTTTTCGAGCCAGTGTAT 300
QY 301 GGAATTGTAATGAAAATTTGGTGAGCGCTTTTAGTCAGAGAGTGCAGGTTCCAGAGGCT 360
Db 301 GGAATTGTAATGAAAATTTGGTGAGCGCTTTTAGTCAGAGAGTGCAGGTTCCAGAGGCT 360
QY 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAGAATGTTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAGAATGTTTCACTCAGAGATGTACAGTTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTTGAA 480
QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATGCGGCTTCGATGGATAGCAGATAGAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAAGCAGATGCGGCTTCGATGGATAGCAGATAGAAA 540
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTTCTGCTGTAAGAGAGGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTTCTGCTGTAAGAGAGGTTTCTTCTCAGGA 600
QY 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGGTTTATGCCAGGACTCACTTTTTC 660
Db 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGGTTTATGCCAGGACTCACTTTTTC 660
QY 661 AATGAACCTCATCAGCAGATGAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 661 AATGAACCTCATCAGCAGATGAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
QY 721 TACTTAGTAAATAAGCTTTCAGAAAGAGGTCAGGAGATCACTTGTGATGCTGCAAA 780
Db 721 TACTTAGTAAATAAGCTTTCAGAAAGAGGTCAGGAGATCACTTGTGATGCTGCAAA 780
QY 781 ATTGAGCAGAGTTTTAAACAGAAAGCTTGCAGAGTGGCTCATTTGGAATGAATTCATT 840
Db 781 ATTGAGCAGAGTTTTAAACAGAAAGCTTGCAGAGTGGCTCATTTGGAATGAATTCATT 840
QY 841 TTGATGAAACAGTACATGAGTTTGTAGCTGACAGATTACTTGTGAACTTGGAACTCTCA 900
Db 841 TTGATGAAACAGTACATGAGTTTGTAGCTGACAGATTACTTGTGAACTTGGAACTCTCA 900
QY 901 AGGTTTTTCAGCAGAAATCCTTTTGAATTTATGGAACATTTCTTTAGAGGAAA 960
Db 901 AGGTTTTTCAGCAGAAATCCTTTTGAATTTATGGAACATTTCTTTAGAGGAAA 960
QY 961 ACAAAATTTCTTTGAGAAAAGGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020
Db 961 ACAAAATTTCTTTGAGAAAAGGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020
QY 1021 ACAGATAACGTTCTTCACTTGGATGAGATTTT 1053
Db 1021 ACAGATAACGTTCTTCACTTGGATGAGATTTT 1053

Db	661	AATGAAC	TCA	TAC	GAC	GAG	ATGA	AGG	ACTT	CAC	TGT	GAC	TTT	GCT	TGC	TGAT	TTCC	NA	720							
Qy	721	TACTTAG	TAA	AT	AG	CC	CTT	CAG	AA	AG	GGT	CAG	G	GAG	AT	CA	T	TT	GT	GT	CA	AAA	780			
Db	721	TACTTAG	TAA	AT	AG	CC	CTT	CAG	AA	AG	GGT	CAG	G	GAG	AT	CA	T	TT	GT	GT	CA	AAA	780			
Qy	781	ATTGAG	CAG	AG	GT	TTT	TAA	CAG	AA	GC	CTT	G	CC	AG	TTG	GC	CT	CA	T	TT	GA	AT	GC	ATT	840	
Db	781	ATTGAG	CAG	AG	GT	TTT	TAA	CAG	AA	GC	CTT	G	CC	AG	TTG	GC	CT	CA	T	TT	GA	AT	GC	ATT	840	
Qy	841	TTGATGA	AC	AG	TAC	AT	TG	AT	TT	G	T	AG	CT	C	A	G	AT	T	TT	G	G	A	AT	TT	CA	900
Db	841	TTGATGA	AC	AG	TAC	AT	TG	AT	TT	G	T	AG	CT	C	A	G	AT	T	TT	G	G	A	AT	TT	CA	900
Qy	901	AAGGT	TTT	T	CAG	G	CA	GA	AA	A	T	T	C	T	T	T	T	T	T	T	T	T	T	T	T	960
Db	901	AAGGT	TTT	T	CAG	G	CA	GA	AA	A	T	T	C	T	T	T	T	T	T	T	T	T	T	T	T	960
Qy	961	ACAA	AT	T	T	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	1020
Db	961	ACAA	AT	T	T	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	1020
Qy	1021	ACAG	ATA	AC	GT	CT	T	C	AC	CT	T	T	G	AT	GC	AG	AT	T	T	T	T	T	T	T	T	1053
Db	1021	ACAG	ATA	AC	GT	CT	T	C	AC	CT	T	T	G	AT	GC	AG	AT	T	T	T	T	T	T	T	T	1053
RESULT 3	BD093077	LOCUS	BD093077	Novel protein and its DNA.	1053 bp	DNA	linear	PAT 27-AUG-2002																		
DEFINITION	BD093077	Novel protein and its DNA.																								
ACCESSION	BD093077	Novel protein and its DNA.																								
VERSION	BD093077.1	GI:22638665																								
KEYWORDS	WO 0100799-A/1.																									
SOURCE	Homo sapiens																									
ORGANISM	Homo sapiens																									
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.																									
AUTHORS	Nakamura,Y., Arakawa,H. and Tanaka,H.																									
TITLE	Novel protein and its DNA																									
JOURNAL	Patent: WO 0100799-A 1 04-JAN-2001; TAKEDA CHEMICAL INDUSTRIES LTD, YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA																									
COMMENT	OS Homo sapiens (human) PN WO 0100799-A/1 PD 04-JAN-2001 PF 27-JUN-2000 WO 2000JP004189 PR 28-JUN-1999 JP 99P 181131,06-JUL-1999 JP 99P 192391 PR 21-JAN-2000 JP 00P 017770 PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC C12N9/04, C12N15/53, C12N1/15, C12N1/19, C12N5/10 PC C12P21/02, A61K38/44, PC A61K48/00, C07K16/40, G01N33/53, C12Q1/26 CC																									
FEATURES	source	1..1053																								
ORIGIN	Query Match 100.0%; Score 1053; DB 6; Length 1053; Best Local Similarity 100.0%; Pred. No. 1.2e-255; Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;																									
Qy	1	ATGGGCGACCGGAAAGCGCGGAGCGCGCGGCTGGATCAGGATCAGAGATCATCTTCA	60																							
Db	1	ATGGGCGACCGGAAAGCGCGGAGCGCGCGGCTGGATCAGGATCAGAGATCATCTTCA	60																							
Qy	61	GACACCAACGAAAGTGAATAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTCTTCGC	120																							

```

VERSION      CQ714252.1  GI:42275109
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
REFERENCE    1
AUTHORS      Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE        Kites, such as nucleic acid arrays, comprising a majority of
              humanexons or transcripts, for detecting expression and other uses
              thereof
JOURNAL      Patent: WO 02068579-A 186 06-SEP-2002;
              PE Corporation (NY) (US)
FEATURES     source
              1..1056
              /organism="Homo sapiens"
              /mol_type="unassigned DNA"
              /db_xref="taxon:9606"
ORIGIN
Query Match      100.0%; Score 1053; DB 6; Length 1056;
Best Local Similarity 100.0%; Pred. No. 1.2e-255;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
DB 1 ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
QY 61 GACACCAACGAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
DB 61 GACACCAACGAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
QY 121 CGGTTGTGTCATCTTCCATCCAGTACCTGATATTTGGAAATGTATAAAGGACAG 180
DB 121 CGGTTGTGTCATCTTCCATCCAGTACCTGATATTTGGAAATGTATAAAGGACAG 180
QY 181 GCTTCCTCTGACAGCAGAGAGGTCGACTTATCAAGGATCTCCCTCACTGGAACAAG 240
DB 181 GCTTCCTCTGACAGCAGAGAGGTCGACTTATCAAGGATCTCCCTCACTGGAACAAG 240
QY 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCAATCTTAGCCCTTTTTCGAGCCAGTGAT 300
DB 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCAATCTTAGCCCTTTTTCGAGCCAGTGAT 300
QY 301 GGAATGTATAAATGAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
DB 301 GGAATGTATAAATGAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
QY 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAATGTTTCACTCAGAGATGTACAGTTTG 420
DB 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAATGTTTCACTCAGAGATGTACAGTTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
DB 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
QY 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGCCCTTGCATGGATAGAGATAGAAA 540
DB 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGCCCTTGCATGGATAGAGATAGAAA 540
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGCTGTGTAAGAGAGTTTCTTCTCAGGA 600
DB 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGCTGTGTAAGAGAGTTTCTTCTCAGGA 600
QY 601 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGCTTTATGCCAGGACTCACTTTTTC 660
DB 601 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGCTTTATGCCAGGACTCACTTTTTC 660
QY 661 AATGAATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTGCTGATGTTCCAA 720
DB 661 AATGAATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTGCTGATGTTCCAA 720
QY 721 TACTTAGTAAATAAGCCTTCAGNAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAA 780
Db 721 TACTTAGTAAATAAGCCTTCAGNAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAA 780
QY 781 ATTGACGAGAGTTTAAACAGAGCCTTCAGAGTTCGCTCATTTGGAAATGAATTCATT 840
Db 781 ATTGACGAGAGTTTAAACAGAGCCTTCAGAGTTCGCTCATTTGGAAATGAATTCATT 840
QY 841 TTGATGAAACAGTACATTTAGTGTGCTCACAGATTTACTTTGTGGAACCTTTGGATTCTCA 900
Db 841 TTGATGAAACAGTACATTTAGTGTGCTCACAGATTTACTTTGTGGAACCTTTGGATTCTCA 900
QY 901 AAGGTTTTTTCAGGAGAAAAATCCTTTTGAATTTTATGGAACCAATTTCTTTAGAGGAAA 960
Db 901 AAGGTTTTTTCAGGAGAAAAATCCTTTTGAATTTTATGGAACCAATTTCTTTAGAGGAAA 960
QY 961 ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTTGCGAGTTATGCGCAGAAC 1020
Db 961 ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTTGCGAGTTATGCGCAGAAC 1020
QY 1021 ACAGATAACGCTTTCACCTTGGATGACAGATTTT 1053
Db 1021 ACAGATAACGCTTTCACCTTGGATGACAGATTTT 1053
RESULT 5
AB166671 1056 bp mRNA linear PRI 09-MAR-2004
LOCUS     Homo sapiens p53R2 mRNA for p53-inducible ribonucleotide reductase
DEFINITION small subunit 2, complete cds.
ACCESSION AB166671
VERSION    AB166671.1 GI:45259568
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
REFERENCE  1
AUTHORS     Ugai, H. and Yokoyama, K.K.
TITLE       Homo sapiens p53R2 mRNA for p53-inducible ribonucleotide reductase
JOURNAL     Published Only in Database (2004)
REFERENCE   2 (bases 1 to 1056)
AUTHORS     Ugai, H. and Yokoyama, K.K.
TITLE       Direct Submission
JOURNAL     Submitted (05-MAR-2004) Hideyo Ugai, RIKEN BioResource Center, Gene
              Engineering Division; 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074,
              Japan (E-mail:ugai@rtc.riken.go.jp, Tel:81-29-836-3612,
              Fax:81-29-836-9120)
FEATURES    Location/Qualifiers
              1..1056
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="p53R2.C19"
              /cell_line="Saos-2"
              /cell_type="Osteogenic sarcoma"
              1..1056
              /gene="p53R2"
              1..1056
              /gene="p53R2"
              /codon_start=1
              /product="p53-inducible ribonucleotide reductase small
              subunit 2"
              /protein_id="BAD12267.1"
              /db_xref="GI:45259569"
              /translation="MGDPERPEAAGLDQDERSSSDTNESEIKSNEEPLIKSRRFVI
              FPQYDPDIWKMYKQAOASFTABEVDLSKDLPHWNKLKADKVFISHILAFPAASDGI
              VNEULVERFSQEVQVPEARCFYQQLIENVHSEMYSLIDTVIISDPKKREFLEFNALIE
              TMPYVKKADWALRWIADRSKTFGRVVAFAAAGVGFSSGFAALFWLKKRGLMGLFT
              FSNELISRDEGLCHDFCLMPQYLVNKPSEERVREIIVDAVKIEQEFLEALPVGILG
              MNCITLMKQYIEFVADRLLELVELGFSKVPQAEFPFDMENISLEGKTNFFEKRVSEYQRF
              AVMAETDNDVFTLDAEP"
ORIGIN

```

Query Match 100.0%; Score 1053; DB 9; Length 1056;
Best Local Similarity 100.0%; Pred. No. 1.2e-255;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGAGATGAGATCATCTTCA 60
DB 1 ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGAGATGAGATCATCTTCA 60

QY 61 GACACCAACGAAGTGAATTAAGTCAATGAAGCCCTCTTAAGAAGAGTCTTCGC 120
DB 61 GACACCAACGAAGTGAATTAAGTCAATGAAGCCCTCTTAAGAAGAGTCTTCGC 120

QY 121 CGGTTTGTCTATCTTTCCATCCAGTACCCTCTGATATTTGGAAATGTATAAACAGGCACAG 180
DB 121 CGGTTTGTCTATCTTTCCATCCAGTACCCTCTGATATTTGGAAATGTATAAACAGGCACAG 180

QY 181 GCTTCTCTTGGACAGCAGAGAGTGCATCTATCAAGGATCTCCCTCACTGGAAACAG 240
DB 181 GCTTCTCTTGGACAGCAGAGAGTGCATCTATCAAGGATCTCCCTCACTGGAAACAG 240

QY 241 CTTAAGCAGATCAGAGTACTTCTCTCTCATCTCTCATCTTACCTTTTTCGAGCCAGTGAT 300
DB 241 CTTAAGCAGATCAGAGTACTTCTCTCTCATCTCTCATCTTACCTTTTTCGAGCCAGTGAT 300

QY 301 GGAAATGTAAATGAATAATTTGGTGGAGCGCTTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
DB 301 GGAAATGTAAATGAATAATTTGGTGGAGCGCTTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360

QY 361 CGCTGTTTCTATGGCTTTCAAATCTCATCGAGATGTTCACTCAGAGATGTACAGTTTG 420
DB 361 CGCTGTTTCTATGGCTTTCAAATCTCATCGAGATGTTCACTCAGAGATGTACAGTTTG 420

QY 421 CTGATAGACACTTATCATGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
DB 421 CTGATAGACACTTATCATGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480

QY 481 ACCATGCCCTATGTTAAGAAAGACAGATTCGGCTTCGGATGGATAGAGTAAAGAA 540
DB 481 ACCATGCCCTATGTTAAGAAAGACAGATTCGGCTTCGGATGGATAGAGTAAAGAA 540

QY 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTGTAAGAGGAGTCTTCTTCAGGA 600
DB 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTGTAAGAGGAGTCTTCTTCAGGA 600

QY 601 TCTTTTGTGCTATATCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
DB 601 TCTTTTGTGCTATATCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660

QY 661 AATGAACCTCATCAGCAGATGAAGACTTCTACTGTGACTTTGCTTGCCTGATGTTCCAA 720
DB 661 AATGAACCTCATCAGCAGATGAAGAGTCTTCTACTGTGACTTTGCTTGCCTGATGTTCCAA 720

QY 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGTCATGCTGTCAAA 780
DB 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGTCATGCTGTCAAA 780

QY 781 ATTGAGCAGGAGTTTAAACAGAGCCTTTCAGATTTGGCTCATTTGGAATGAATTGCATT 840
DB 781 ATTGAGCAGGAGTTTAAACAGAGCCTTTCAGATTTGGCTCATTTGGAATGAATTGCATT 840

QY 841 TTGATGAACAGTACATTTAGTTTGTAGCTGACAGATTAATCTTGTGGAACCTTGATTTCTCA 900
DB 841 TTGATGAACAGTACATTTAGTTTGTAGCTGACAGATTAATCTTGTGGAACCTTGATTTCTCA 900

QY 901 AAGGTTTTTTCAGGCAGAAAAATCCTTTTGAATTTATGGAAAAATTTCTTTAGAGGAAA 960
DB 901 AAGGTTTTTTCAGGCAGAAAAATCCTTTTGAATTTATGGAAAAATTTCTTTAGAGGAAA 960

QY 961 ACAAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGAGTTATGGCAGAAACC 1020
DB 961 ACAAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGAGTTATGGCAGAAACC 1020

QY 1021 ACAGATAACGTCCTTCACTTGGATGCAGATTTT 1053
DB 1021 ACAGATAACGTCCTTCACTTGGATGCAGATTTT 1053

RESULT 6
LOCUS AR454869 4955 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 3 from patent US 6682917.
ACCESSION AR454869
VERSION AR454869.1 GI:42688824
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4955)
AUTHORS Nakamura,Y., Arakawa,H. and Tanaka,H.
TITLE Protein having a ribonucleotide Reductase activity and a DNA thereof
JOURNAL Patent: US 6682917-A 3 27-JAN-2004;
FEATURES Location/Qualifiers
 source
 1..4955
 /mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 1053; DB 6; Length 4955;
Best Local Similarity 100.0%; Pred. No. 1.1e-255;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAGCGCGGCTGGATCAGGATGAGATCATCTTCA 60
DB 245 ATGGGCGACCCGGAAGCGCGGCTGGATCAGGATGAGATCATCTTCA 304

QY 61 GACACCAACGAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTCTTCGC 120
DB 305 GACACCAACGAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTCTTCGC 364

QY 121 CGGTTTGTCTATCTTTCCATCCAGTACCTGTATATTTGGAAATGTATAAACAGGCACAG 180
DB 365 CGGTTTGTCTATCTTTCCATCCAGTACCTGTATATTTGGAAATGTATAAACAGGCACAG 424

QY 181 GCTTCTCTTGGACAGCAGAGAGTGCAGTTCATCAAGAGTCTCCCTCACTGGAAACAG 240
DB 425 GCTTCTCTTGGACAGCAGAGAGTGCAGTTCATCAAGAGTCTCCCTCACTGGAAACAG 484

QY 241 CTTAAGCAGATCAGAGTACTTCTCATCTCACAATCTTAGCCTTTTTCGAGCCAGTGAT 300
DB 485 CTTAAGCAGATCAGAGTACTTCTCATCTCACAATCTTAGCCTTTTTCGAGCCAGTGAT 544

QY 301 GGAATTTGAATGAATAATTTGGTGGAGCGCTTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
DB 545 GGAATTTGAATGAATAATTTGGTGGAGCGCTTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 604

QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATGAGNATGTTCACTCAGAGATGTACAGTTTG 420
DB 605 CGCTGTTTCTATGGCTTTCAAAATCTCATGAGNATGTTCACTCAGAGATGTACAGTTTG 664

QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
DB 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 724

QY 481 ACCATGCCCTATGTTAAGAAAGACAGATTTGGCCCTTTCGATGGATAGCAGATAGAAAA 540
DB 725 ACCATGCCCTATGTTAAGAAAGACAGATTTGGCCCTTTCGATGGATAGCAGATAGAAAA 784

QY 541 TCTACTTTTGGGAAAGAGTGTGGCTTTTCTGCTGTAGAGAGTCTTCTTCTCAGGA 600
DB 785 TCTACTTTTGGGAAAGAGTGTGGCTTTTCTGCTGTAGAGAGTCTTCTTCTCAGGA 844

QY 601 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGGTCCTTATGCCAGGACTCACATTTTTC 660
DB 845 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGGTCCTTATGCCAGGACTCACATTTTTC 904

Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
QY	1 ATGGGCGACCCGGAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db	245 ATGGGCGACCCGGAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
QY	61 GACACCAAGTGAATGAAGTCAATGAAGCCCACTCTTAAGAAAGAGTTCTCGC 120
Db	305 GACACCAAGTGAATGAAGTCAATGAAGCCCACTCTTAAGAAAGAGTTCTCGC 364
QY	121 CGTTTTGTCTCTTTTCCAATCCAGTACCTCTGATATTTGGAAATATGTAACAAGGCACAG 180
Db	365 CGTTTTGTCTCTTTTCCAATCCAGTACCTCTGATATTTGGAAATATGTAACAAGGCACAG 424
QY	181 GCTTCCTCTTCGACAGCAGAGAGGTGCACTTATCAAAGGATCTCCCTCACTCGAAACAAG 240
Db	425 GCTTCCTCTTCGACAGCAGAGAGGTGCACTTATCAAAGGATCTCCCTCACTCGAAACAAG 484
QY	241 CTTAAAGCAGATGAGAAGTACTTCATCTCTCACTCTTAGCCCTTTTTCGAGCCAGTGAT 300
Db	485 CTTAAAGCAGATGAGAAGTACTTCATCTCTCACTCTTAGCCCTTTTTCGAGCCAGTGAT 544
QY	301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
Db	545 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 604
QY	361 CGCTGTTTCTATGGCTTTTCAAAATTTCTCATCGAAATGTTTCACTCAGAGATGTACAGTTTG 420
Db	605 CGCTGTTTCTATGGCTTTTCAAAATTTCTCATCGAAATGTTTCACTCAGAGATGTACAGTTTG 664
QY	421 CTGATAGACCTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTTGAA 480
Db	665 CTGATAGACCTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTTGAA 724
QY	481 ACCATGCCCTATGTTTAAGAAAGAGAGATTTGGCCCTTCGGATGGATAGCAGATAGAAAA 540
Db	725 ACCATGCCCTATGTTTAAGAAAGAGAGATTTGGCCCTTCGGATGGATAGCAGATAGAAAA 784
QY	541 TCTACTTTTTGGGAAAGAGTGGTGGCTTTTGCTGCTGTAGAAAGAGTTTTTCTTCTCAGGA 600
Db	785 TCTACTTTTTGGGAAAGAGTGGTGGCTTTTGCTGCTGTAGAAAGAGTTTTTCTTCTCAGGA 844
QY	601 TCTTTTGTGCTATATTTCTGGCTTAAAGAAAGAGAGTCTTATGCGAGACTCACTTTTTTCC 660
Db	845 TCTTTTGTGCTATATTTCTGGCTTAAAGAAAGAGAGTCTTATGCGAGACTCACTTTTTTCC 904
QY	661 AATGAATCATCAGCAGAGATGAAGACTTCACTGTGACTTTTCTGCTGCTGATGTTTCCAA 720
Db	905 AATGAATCATCAGCAGAGATGAAGACTTCACTGTGACTTTTCTGCTGCTGATGTTTCCAA 964
QY	721 TACTTAGTAAATAAGCTTTTCAAGAAAGAGGTTCAGGAGATCAATTTGATGCTGTCAAA 780
Db	965 TACTTAGTAAATAAGCTTTTCAAGAAAGAGGTTCAGGAGATCAATTTGATGCTGTCAAA 1024
QY	781 ATTGAGCAGAGTTTTTAAACAGAGCTTGCACAGTTGGCTCATTTGGAATGAATTCGATT 840
Db	1025 ATTGAGCAGAGTTTTTAAACAGAGCTTGCACAGTTGGCTCATTTGGAATGAATTCGATT 1084
QY	841 TTGATGAACAGTACATTTGAGTTTGTAGCTGCAGAGATTTACTTTGGAACTTGGATTCTCA 900
Db	1085 TTGATGAACAGTACATTTGAGTTTGTAGCTGCAGAGATTTACTTTGGAACTTGGATTCTCA 1144
QY	901 AAGGTTTTTTCAGCAGAAAATCCTTTTGATTTTATGAAAACATTTCTTTAGAGGAAA 960
Db	1145 AAGGTTTTTTCAGCAGAAAATCCTTTTGATTTTATGAAAACATTTCTTTAGAGGAAA 1204
QY	961 ACAAATTTCTTCAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGGTTATGGCAGAAACC 1020
Db	1205 ACAAATTTCTTCAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGGTTATGGCAGAAACC 1264
QY	1021 ACAGATAACGCTTTCACCTTTGGATGCAGATTTT 1053
Db	1265 ACAGATAACGCTTTCACCTTTGGATGCAGATTTT 1297
RESULT 7	
LOCUS	BD064765 4955 bp DNA linear PAT 27-AUG-2002
DEFINITION	Novel protein and DNA thereof.
ACCESSION	BD064765
VERSION	BD064765.1 GI:22610368
KEYWORDS	JP 2001269184-A/2.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1. (bases 1 to 4955)
TITLE	Nakamura, Y., Arakawa, H. and Tanaka, H.
JOURNAL	Novel protein and DNA thereof
COMMENT	Patent: JP 2001269184-A 2 02-OCT-2001; YUSUKE NAKAMURA, TAKEDA CHEMICAL INDUSTRIES LTD OS Homo sapiens (human) PN JP 2001269184-A/2 PD 02-OCT-2001 PP 27-JUN-2000 JP 2000192401 PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC C12N15/09, A61K31/711, A61K38/00, A61K38/43, A61K45/00, A61K48/00, PC A61P35/00, PC A61P43/00, C07K16/32, C07K16/40, C12N1/15, C12N1/19, C12N1/21 PC C12N5/10, C12N9/02, PC C12P21/02, C12P21/08, G01N33/15, G01N33/50, G01N33/53// (C12N15/09, PC C12R1:91), PC (C12N15/09, C12R1:19), (C12N9/02, C12R1:91), (C12P21/02, C12R1:91), PC (C12P21/08, C12R1:91), C12N15/00, A61K37/02, A61K37/48, C12N5/00, PC (C12N15/00, C12R1:91), (C12N15/00, C12R1:19) CC Novel protein and DNA thereof FH Key Location/Qualifiers FT source 1. .4955 FT /organism='Homo sapiens (human)'. FT Location/Qualifiers 1. .4955 /organism='Homo sapiens' /mol_type='genomic DNA' /db_xref='taxon:9606'
FEATURES	source
ORIGIN	Query Match 100.0%; Score 1053; DB 6; Length 4955; Best Local Similarity 100.0%; Pred. No. 1.1e-255;

Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
QY	1 ATGGGCGACCGCGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db	245 ATGGGCGACCGCGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
QY	61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCCTAAGAAAGAGTTCTCGC 120
Db	305 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCCTAAGAAAGAGTTCTCGC 364
QY	121 CGGTTTGTGATCTTTTCAATCCAGTACCTTGATATTTGGAAAAATGTATAAAGCAGCAG 180
Db	365 CGGTTTGTGATCTTTTCAATCCAGTACCTTGATATTTGGAAAAATGTATAAAGCAGCAG 424
QY	181 GCCTTCCTTCTGGACAGCAGAGAGGTGCACTTATCAAAAGGATCTCCCTCACTCGAAACAAG 240
Db	425 GCCTTCCTTCTGGACAGCAGAGAGGTGCACTTATCAAAAGGATCTCCCTCACTCGAAACAAG 484
QY	241 CTTAAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTAGCCCTTTTTCAGCGCAGTGAT 300
Db	485 CTTAAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTAGCCCTTTTTCAGCGCAGTGAT 544
QY	301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
Db	545 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 604
QY	361 CGCTGTTTCTATGGCTTTCAAAATTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 420
Db	605 CGCTGTTTCTATGGCTTTCAAAATTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 664
QY	421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAAATGCAATTTGAA 480
Db	665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAAATGCAATTTGAA 724
QY	481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGCATGATGATGATGATGATGAT 540
Db	725 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGCATGATGATGATGATGATGAT 784
QY	541 TCTACTTTTGGGNAAGAGTGGTGGCTTTTGTGCTGTAGAGAGGATTTTCTTCTCAGGA 600
Db	785 TCTACTTTTGGGNAAGAGTGGTGGCTTTTGTGCTGTAGAGAGGATTTTCTTCTCAGGA 844
QY	601 TCTTTTCTGCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 660
Db	845 TCTTTTCTGCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 904
QY	661 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGATGCTTGTGCTGATGTTCCAA 720
Db	905 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGATGCTTGTGCTGATGTTCCAA 964
QY	721 TACTTAGTAATAAGCCTTCAGAAAGAGGTTCAGGAGATCATTTGTGATGCTGTCAAA 780
Db	965 TACTTAGTAATAAGCCTTCAGAAAGAGGTTCAGGAGATCATTTGTGATGCTGTCAAA 1024
QY	781 ATTGAGCAGAGTATTTTAAACAGAGCTTGCAGTGGCCCTCATTTGGAATGAATTCGATT 840
Db	1025 ATTGAGCAGAGTATTTTAAACAGAGCTTGCAGTGGCCCTCATTTGGAATGAATTCGATT 1084
QY	841 TTGATGAACAGTACATTTGATGTTGTAGCTGCAGAGATTACTTGTGGAACCTTGGATTCTCA 900
Db	1085 TTGATGAACAGTACATTTGATGTTGTAGCTGCAGAGATTACTTGTGGAACCTTGGATTCTCA 1144
QY	901 AAGGTTTTTTCAGGCAGAAAAATCCTTTTGTATTTTATGAAAAATTTCTTTTAGAAGAAAA 960
Db	1145 AAGGTTTTTTCAGGCAGAAAAATCCTTTTGTATTTTATGAAAAATTTCTTTTAGAAGAAAA 1204
QY	961 ACAAATTTCTTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGGTTATGGCAGAAACC 1020
Db	1205 ACAAATTTCTTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGGTTATGGCAGAAACC 1264
QY	1021 ACAGATAACGCTCTTACCTTGGATGCAGATTTT 1053
Db	1265 ACAGATAACGCTCTTACCTTGGATGCAGATTTT 1297

RESULT 8	BD093078	Novel protein and its DNA.	4955 bp	DNA	linear	PAT 27-AUG-2002
LOCUS	BD093078					
DEFINITION	BD093078					
ACCESSION	BD093078					
VERSION	BD093078.1	GI:22638666				
KEYWORDS	WO 0100799-A/2.					
SOURCE	WO 0100799-A/2.					
ORGANISM	Homo sapiens					
REFERENCE	1 (bases 1 to 4955)					
AUTHORS	Nakamura,Y., Arakawa,H. and Tanaka,H.					
TITLE	Novel protein and its DNA					
JOURNAL	Patent: WO 0100799-A 2 04-JAN-2001; TAKEDA CHEMICAL INDUSTRIES LTD, YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA					
COMMENT	OS Homo sapiens (human) PN WO 0100799-A/2 PD 04-JAN-2001 PF 27-JUN-2000 WO 2000JP004189 PR 28-JUN-1999 JP 99P 181133,06-JUL-1999 JP 99P 192391 PR 21-JAN-2000 JP 00P 017770 PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC C12N9/04, C12N15/53, C12N1/15, C12N1/19, C12N1/21, C12N5/10 PC C12P21/02, A61K38/44, PC A61K48/00, C07K16/40, G01N33/53, C12Q1/26 CC					
PH	Key	Location/Qualifiers.				
FEATURES	source	1..4955				
		/organism="Homo sapiens"				
		/mol_type="genomic DNA"				
		/db_xref="taxon:9606"				
ORIGIN						
	Query Match	100.0%;	Score 1053;	DB 6;	Length 4955;	
	Best Local Similarity	100.0%;	Pred. No. 1.1e-255;			
	Matches 1053;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	1	ATGGGGACCCGAAAGCGCGGAGCGCGGCTGGATCAGGATCAGAGATCATCTTCA	60			
DB	245	ATGGGACCCGAAAGCGCGGAGCGCGGCTGGATCAGGATCAGAGATCATCTTCA	304			
QY	61	GACACCAAGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCCTGC	120			
DB	305	GACACCAAGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCCTGC	364			
QY	121	CGGTTTGTGATCTTTCCATCCAGTACCTCGATATTGGAAATGTATAACAGGCACAG	180			
DB	365	CGGTTTGTGATCTTTCCATCCAGTACCTCGATATTGGAAATGTATAACAGGCACAG	424			
QY	181	GCTTCTCTTCGGACAGCAGAAAGAGTTCGACTTATCAAAAGGATCTCCCTCACTGGAAACAAG	240			
DB	425	GCTTCTCTTCGGACAGCAGAAAGAGTTCGACTTATCAAAAGGATCTCCCTCACTGGAAACAAG	484			
QY	241	CTTAAAGCAGATGAGAAGTACTTCACTCTCACTCTTACATCTTACAGCCCTTTTTCAGCCAGTGAT	300			
DB	485	CTTAAAGCAGATGAGAAGTACTTCACTCTCACTCTTACATCTTACAGCCCTTTTTCAGCCAGTGAT	544			
QY	301	GGAAATGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT	360			
DB	545	GGAAATGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT	604			
QY	361	CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAAATGTTCACTCAGAGATGTACAGTTTG	420			
DB	605	CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAAATGTTCACTCAGAGATGTACAGTTTG	664			
QY	421	CTGATGACCTTACATCAGAGATCCCAAGAAAGGGAATTTTTTAAATGCAATGAA	480			
DB	665	CTGATGACCTTACATCAGAGATCCCAAGAAAGGGAATTTTTTAAATGCAATGAA	724			


```
Db 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTGTCAGAGGCT 360
QY 361 CGCTGTTTCTATGGCTTTTCAATTTCTCATCGAGAATGTTTCACTCAGAGATGTACAGTTG 420
Db 361 CGCTGTTTCTATGGCTTTTCAATTTCTCATCGAGAATGTTTCACTCAGAGATGTACAGTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTAAATGCAATTTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTAAATGCAATTTGAA 480
QY 481 ACATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGCATGAGATAGCAGATAGAAA 540
Db 481 ACATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGCATGAGATAGCAGATAGAAA 540
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAGAGATTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAGAGATTTTCTTCTCAGGA 600
QY 601 TCTTTTGTCTATATTTCTGGCTTAAAGAAAGAGAGGCTTTATGCGAGAGCTCACTTTTTC 660
Db 601 TCTTTTGTCTATATTTCTGGCTTAAAGAAAGAGAGGCTTTATGCGAGAGCTCACTTTTTC 660
QY 661 AATGAATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 661 AATGAATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
QY 721 TACTTAGTAAATAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 780
Db 721 TACTTAGTAAATAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 780
QY 781 ATTGAGCAGAGTTTTAAACAGAGCCTTGCAGAGCTTGGCTCATTTGGAATGCAAT 840
Db 781 ATTGAGCAGAGTTTTAAACAGAGCCTTGCAGAGCTTGGCTCATTTGGAATGCAAT 840
QY 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATTTGTGAACTTGTGATTTCA 900
Db 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATTTGTGAACTTGTGATTTCA 900
QY 901 AAGTTTTTTCAGCAGAGAAATCCTTTGATTTTATGGAAGAAATTTCTTTAGAGGAA 960
Db 901 AAGTTTTTTCAGCAGAGAAATCCTTTGATTTTATGGAAGAAATTTCTTTAGAGGAA 960
QY 961 ACAATTTCTTTGAGAAACAGTTTCAGAGATATCAGCTTTTGCAGTTATGCGAGAAACC 1020
Db 961 ACAATTTCTTTGAGAAACAGTTTCAGAGATATCAGCTTTTGCAGTTATGCGAGAAACC 1020
QY 1021 ACAGATAAGCTCTTCACTTGGATGAGATTTT 1053
Db 1021 ACAGATAAGCTCTTCACTTGGATGAGATTTT 1053

RESULT 11
LOCUS BD064773 1053 bp DNA linear PAT 27-AUG-2002
DEFINITION Novel protein and DNA thereof.
ACCESSION BD064773
VERSION BD064773.1 GI:22610376
KEYWORDS JP 2001269184-A/10.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1053)
AUTHORS Nakamura, Y., Arakawa, H. and Tanaka, H.
TITLE Novel protein and DNA thereof
JOURNAL Patent: JP 2001269184-A 10 02-OCT-2001;
YUSUKE NAKAMURA, TAKEDA CHEMICAL INDUSTRIES LTD
COMMENT OS Homo sapiens (human)
PN JP 2001269184-A/10
PD 02-OCT-2001
PF 27-JUN-2000 JP 2000192401
PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
```

```
FEATURES
source 1..1053
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 99.8%; Score 1051.4; DB 6; Length 1053;
Best Local Similarity 99.9%; Pred. No. 3e-255;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGGACCCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 ATGGCGGACCCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
QY 61 GACACCAAGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTGC 120
Db 61 GACACCAAGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTGC 120
QY 121 CGTTTGTCTATCTTCCATCCAGTACCTGATATTTGGAATAATGTAACAGGCACAG 180
Db 121 CGTTTGTCTATCTTCCATCCAGTACCTGATATTTGGAATAATGTAACAGGCACAG 180
QY 181 GCTTCTCTTGGACAGAGAGGTCGACTTATCAAGAGATCTCCCTCAGTGGAAACAG 240
Db 181 GCTTCTCTTGGACAGAGAGGTCGACTTATCAAGAGATCTCCCTCAGTGGAAACAG 240
QY 241 CTTAAAGCAGATGAGAGTACTTCTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTGT 300
Db 241 CTTAAAGCAGATGAGAGTACTTCTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTGT 300
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGAGTGCAAGTTCAGAGCT 360
Db 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGAGTGCAAGTTCAGAGCT 360
QY 361 CGCTGTTTCTATGGCTTTCAATTTCTCATCGAGAATGTTTCACTCAGAGATGTACAGTTG 420
Db 361 CGCTGTTTCTATGGCTTTCAATTTCTCATCGAGAATGTTTCACTCAGAGATGTACAGTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTAAATGCAATTTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTAAATGCAATTTGAA 480
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGCATGAGATAGCAGATAGAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGCATGAGATAGCAGATAGAAA 540
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAGAGATTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAGAGATTTTCTTCTCAGGA 600
QY 601 TCTTTTGTCTATATTTCTGGCTTAAAGAAAGAGAGGCTTTATGCGAGAGCTCACTTTTTC 660
Db 601 TCTTTTGTCTATATTTCTGGCTTAAAGAAAGAGAGGCTTTATGCGAGAGCTCACTTTTTC 660
QY 661 AATGAATCATCAGCAGAGATGAAGAGCTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 661 AATGAATCATCAGCAGAGATGAAGAGCTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
QY 721 TACTTAGTAAATAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 780
```

Db 721 TACTTAGTAAATAAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGTGATGCTGTCAAA 780
 Qy 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTTGGCTCATTTGGAATGAATTCATT 840
 Db 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTTGGCTCATTTGGAATGAATTCATT 840
 Qy 841 TTGATGAACACGTACATTCAGTTTGTAGCTGACAGATTAATCTTGTGAACTTTGGATTCTCA 900
 Db 841 TTGATGAACACGTACATTCAGTTTGTAGCTGACAGATTAATCTTGTGAACTTTGGATTCTCA 900
 Qy 901 AAGGTTTTTTCAGGAGAAAATCCTTTTGAATTTATGAAAACATTTCTTTAGAGGAAAA 960
 Db 901 AAGGTTTTTTCAGGAGAAAATCCTTTTGAATTTATGAAAACATTTCTTTAGAGGAAAA 960
 Qy 961 ACAAAATTTCTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGGATGAGCAAAACC 1020
 Db 961 ACAAAATTTCTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGGATGAGCAAAACC 1020
 Qy 1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053
 Db 1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053

RESULT 12
 BD093086
 LOCUS Novel protein and its DNA. linear PAT 27-AUG-2002
 DEFINITION BD093086
 ACCESSION BD093086
 VERSION BD093086.1 GI:22638674
 KEYWORDS WO 0100799-A/10.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1053)
 AUTHORS Nakamura, Y., Arakawa, H. and Tanaka, H.
 TITLE Novel protein and its DNA
 JOURNAL TAKEDA CHEMICAL INDUSTRIES LTD, YUSUKE NAKAMURA, HIROFUMI ARAKAWA,
 HIROSHI TANAKA
 COMMENT PN WO 0100799-A/10
 PD 04-JAN-2001
 PF 27-JUN-2000 WO 2000JP004189
 PR 28-JUN-1999 JP 99P 181131, 06-JUL-1999 JP 99P 192391 PR
 21-JAN-2000 JP 00P 017770
 PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
 C12N9/04, C12N15/53, C12N1/15, C12N1/19, C12N1/21, C12N5/10 PC
 C12P21/02, A61K38/44,
 PC A61K48/00, C07K16/40, G01N33/53, C12Q1/26
 CC
 FH Key Location/Qualifiers.

FEATURES
 source
 1..1053
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 Query Match 99.8%; Score 1051.4; DB 6; Length 1053;
 Best Local Similarity 99.8%; Pred. No. 3e-255;
 Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 ORIGIN
 Qy 1 ATGGGGGACCCGGAAGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
 Db 1 ATGGGGGACCCGGAAGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
 Qy 61 GACACCAACGAAAGTGAATTAAGTCAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
 Db 61 GACACCAACGAAAGTGAATTAAGTCAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
 Qy 121 CGGTTTGTCAATCTTCCATCCAGTACCCTGTATTTGGAAATGTATAACAGGCACAG 180

Db 121 CGGTTTGTCAATCTTCCATCCAGTACCCTGTATTTGGAAATGTATAACAGGCACAG 180
 Qy 181 GCTTCCTTCTGGACAGCAGAGAGGTGACATTATCAAGAGATCTCCCTCACTGGACAG 240
 Db 181 GCTTCCTTCTGGACAGCAGAGAGGTGACATTATCAAGAGATCTCCCTCACTGGACAG 240
 Qy 241 CTTAAAGCAGATGAGAGTACTTTCATCTCTCACATCTTAGCCTTTTTCAGAGCCAGTGT 300
 Db 241 CTTAAAGCAGATGAGAGTACTTTCATCTCTCACATCTTAGCCTTTTTCAGAGCCAGTGT 300
 Qy 301 GGAATGTAAATGAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
 Db 301 GGAATGTAAATGAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
 Qy 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAATGTTTCACTCAGAGATGTACAGTTG 420
 Db 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAATGTTTCACTCAGAGATGTACAGTTG 420
 Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAAATGCAATTGAA 480
 Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAAATGCAATTGAA 480
 Qy 481 ACCATGCCCTATGTTAAAGAAAAGAGATTTGGCCCTTTGCGATGGATAGCAGATGAAAA 540
 Db 481 ACCATGCCCTATGTTAAAGAAAAGAGATTTGGCCCTTTGCGATGGATAGCAGATGAAAA 540
 Qy 541 TCTACTTTGGGAAAGAGTGGTGGCTTCTGCTGCTGTAGAGAGTCTTCTTCTCAGGA 600
 Db 541 TCTACTTTGGGAAAGAGTGGTGGCTTCTGCTGCTGTAGAGAGTCTTCTTCTCAGGA 600
 Qy 601 TCTTTGCTGCTATATTCTGCTAAAGAAAGAGAGTCTTATGCCAGGAGTCACTTTTCTCC 660
 Db 601 TCTTTGCTGCTATATTCTGCTAAAGAAAGAGAGTCTTATGCCAGGAGTCACTTTTCTCC 660
 Qy 661 AATGAACCTCATCAGCAGAGATGAAGAGCTTCACTGTGACTTTGCTTGCTGATGTTCCAA 720
 Db 661 AATGAACCTCATCAGCAGAGATGAAGAGCTTCACTGTGACTTTGCTTGCTGATGTTCCAA 720
 Qy 721 TACTTAGTAAATAGCCTTCAGAAAGAGGCTGAGGAGATCATTTGTGATGCTGTCAAA 780
 Db 721 TACTTAGTAAATAGCCTTCAGAAAGAGGCTGAGGAGATCATTTGTGATGCTGTCAAA 780
 Qy 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTTGGCCTCATTTGGAATGAATTCATT 840
 Db 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTTGGCCTCATTTGGAATGAATTCATT 840
 Qy 841 TTGATGAACACAGTACATTTGAGTTTGTAGCTGACAGATTAATCTTGTGAACTTTGGATTCTCA 900
 Db 841 TTGATGAACACAGTACATTTGAGTTTGTAGCTGACAGATTAATCTTGTGAACTTTGGATTCTCA 900
 Qy 901 AAGGTTTTTTCAGGAGAAAATCCTTTTGAATTTATGAAAACATTTCTTTAGAGGAAAA 960
 Db 901 AAGGTTTTTTCAGGAGAAAATCCTTTTGAATTTATGAAAACATTTCTTTAGAGGAAAA 960
 Qy 961 ACAAAATTTCTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGGATGAGCAAAACC 1020
 Db 961 ACAAAATTTCTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGGATGAGCAAAACC 1020
 Qy 1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053
 Db 1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053

RESULT 13
 AR454870
 LOCUS Sequence 4 from patent US 6682917. linear PAT 20-FEB-2004
 DEFINITION AR454870
 ACCESSION AR454870
 VERSION AR454870.1 GI:42688825
 KEYWORDS Unknown.
 SOURCE ORGANISM Unknown.

Db	860	TTGATGAAACAGTACATTTGATTTGTAGCTGACAGATTACTTGTGGAACATTGGAATCTCTCA	919
Qy	901	AAGGTTTTTCAGGCAGAAAAATCCTTTTGTATTTATGMAAACATTTCTTTAGAAAGGAAAA	960
Db	920	AAGGTTTTTCAGGCAGAAAAATCCTTTTGTATTTATGMAAACATTTCTTTAGAAAGGAAAA	979
Qy	961	ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC	1020
Db	980	ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC	1039
Qy	1021	ACAGATAACGCTCTTCACCTTGGATGCAGATTTT	1053
Db	1040	ACAGATAACGCTCTTCACCTTGGATGCAGATTTT	1072
RESULT 14			
BD064766			
LOCUS	BD064766	1081 bp	DNA linear PAT 27-AUG-2002
DEFINITION	Novel protein and DNA thereof.		
ACCESSION	BD064766		
VERSION	BD064766.1 GI:22610369		
KEYWORDS	JP 2001269184-A/3.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 1081)		
TITLE	Nakamura, Y., Arakawa, H. and Tanaka, H.		
JOURNAL	Novel protein and DNA thereof		
COMMENT	Patent: JP 2001269184-A 3 02-OCT-2001; YUSUKE NAKAMURA, TAKEDA CHEMICAL INDUSTRIES LTD OS Homo sapiens (human) PN JP 2001269184-A/3 PD 02-OCT-2001 PF 27-JUN-2000 JP 2000192401 PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC C12N15/09, A61K31/711, A61K38/00, A61K38/43, A61K45/00, A61K48/00, PC A61P35/00, PC A61P43/00, C07K16/32, C07K16/40, C12N1/15, C12N1/19, C12N1/21 PC C12N5/10, C12N9/02, PC C12P21/02, C12P21/08, G01N33/15, G01N33/50, G01N33/53/(C12N15/09, C12N1/91), PC (C12N15/09, C12R1:19), (C12N9/02, C12R1:91), (C12P21/02, C12R1:91), PC (C12P21/08, C12R1:91), C12N15/00, A61K37/02, A61K37/48, C12N5/00, PC (C12N15/00, C12R1:91), (C12N15/00, C12R1:19) CC Novel protein and DNA thereof CH Key Location/Qualifiers FH Key Location/Qualifiers FT source 1..1081 FT Location/Qualifiers 1..1081 /organism='Homo sapiens (human)'		
FEATURES	source		
ORIGIN	Query Match 99.8%; Score 1051.4; DB 6; Length 1081; Best Local Similarity 99.9%; Pred. No. 3e-255; Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		
Qy	1	ATGGGGCAGCCCGAAAGCGGGAAGCGCGGGCTGGATCAGATCAGAGTCAATCTTCA	60
Db	20	ATGGGGCAGCCCGAAAGCGGGAAGCGCGGGCTGGATCAGATCAGAGTCAATCTTCA	79
Qy	61	GACACCAAGTGAATGAAGTCAAAATGAAGGCCACTCTCTAAGAAAGAGTTCTCGC	120
Db	80	GACACCAAGTGAATGAATGAAGTCAAAATGAAGGCCACTCTCTAAGAAAGAGTTCTCGC	139
Qy	121	CGGTTTGTCTTTCCCAATCCAGTACCTTGATATTTGGAAAATGTATATAACAGGCACAG	180
Db	140	CGGTTTGTCTTTCCCAATCCAGTACCTTGATATTTGGAAAATGTATATAACAGGCACAG	199

```
QY 181 GCTTCTCTTGTGACAGCAGAGAGGTCGACTTATCAAGAGTCTCCCTCACTGGACACAG 240
Db 200 GCTTCTCTTGTGACAGCAGAGAGGTTGACTTATCAAGAGTCTCCCTCACTGGACACAG 259
QY 241 CTTAAAGCAGATGAGAGTACTTCTCTCACATCTTAGCCCTTTTGTGAGCCAGTGAT 300
Db 260 CTTAAAGCAGATGAGAGTACTTCTCTCACATCTTAGCCCTTTTGTGAGCCAGTGAT 319
QY 301 CGAAATGTAATGAAATTTGTGGAGCGCTTTAGTCAGAGGTGTCAGAGGCT 360
Db 320 GGAATGTAATGAAATTTGTGGAGCGCTTTAGTCAGAGGTGTCAGAGGCT 379
QY 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db 380 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 439
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 440 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 499
QY 481 ACCATGCCCTATGTTAAGAAAGGAGCAGATTTGGGCTTGGATGGATAGCAGATAGAAA 540
Db 500 ACCATGCCCTATGTTAAGAAAGGAGCAGATTTGGGCTTGGATGGATAGCAGATAGAAA 559
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGTCTTCTTCTCAGGA 600
Db 560 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGTCTTCTTCTCAGGA 619
QY 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGCTTTATGCGCAGACTCACTTTTTC 660
Db 620 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGCTTTATGCGCAGACTCACTTTTTC 679
QY 661 AATGAACTCATCAGCAGATGAGAGCTTCACTGTGACTTTGCTTGGCTGATGTTCCAA 720
Db 680 AATGAACTCATCAGCAGATGAGAGCTTCACTGTGACTTTGCTTGGCTGATGTTCCAA 739
QY 721 TACTTAGTAAATAAGCTTTCAAGAAAGGCTCAGGAGATCATTTGTCCTGATGTTCCAA 780
Db 740 TACTTAGTAAATAAGCTTTCAAGAAAGGCTCAGGAGATCATTTGTCCTGATGTTCCAA 799
QY 781 ATTGAGCAGAGTATTTTAAAGAGGCTTTGCCAGTTGGCTCATTTGGAATGAATTCAT 840
Db 800 ATTGAGCAGAGTATTTTAAAGAGGCTTTGCCAGTTGGCTCATTTGGAATGAATTCAT 859
QY 841 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTAAGTGTGAACTTGTCA 900
Db 860 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTAAGTGTGAACTTGTCA 919
QY 901 AAGGTTTTTTCAGCAGAAATCCTTTTGTATTTTATGAAACATTTTCTTTAGAGGAAA 960
Db 920 AAGGTTTTTTCAGCAGAAATCCTTTTGTATTTTATGAAACATTTTCTTTAGAGGAAA 979
QY 961 ACAATTTTCTTTGAGAAACAGTTTTCAGAGTATCAGCGTTTGTGAGTTATGCGAGAAAC 1020
Db 980 ACAATTTTCTTTGAGAAACAGTTTTCAGAGTATCAGCGTTTGTGAGTTATGCGAGAAAC 1039
QY 1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053
Db 1040 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1072
```

```
RESULT 15
BD093079
LOCUS BD093079 1081 bp DNA linear PAT 27-AUG-2002
DEFINITION Novel protein and its DNA.
ACCESSION BD093079
VERSION BD093079.1 GI:22638667
KEYWORDS WO 0100799-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1081)
```

AUTHORS
TITLE
JOURNAL

Nakamura, Y., Arakawa, H. and Tanaka, H.
Novel protein and its DNA
Patent: WO 0100799-A 3 04-JAN-2001;
TAKEDA CHEMICAL INDUSTRIES LTD, YUSUKE NAKAMURA, HIROFUMI ARAKAWA,
HIROSHI TANAKA

COMMENT

OS Homo sapiens (human)
PN WO 0100799-A/3
PD 04-JAN-2001
PF 27-JUN-2000 WO 2000JP004189
PR 28-JUN-1999 JP 99P 181131, 06-JUL-1999 JP 99P 192391 PR
21-JAN-2000 JP 00P 017770
PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
C12N9/04, C12N15/53, C12N1/15, C12N1/19, C12N5/10 PC
C12P21/02, A61K38/44,
PC A61K48/00, C07K16/40, G01N33/53, C12Q1/26
CC

FEATURES

source

1. .1081
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 99.8%; Score 1051.4; DB 6; Length 1081;
Best Local Similarity 99.9%; Pred. No. 3e-255;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 20 ATGGGCGACCCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 79
QY 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCCCTAAGAAAGAGTTCTCGC 120
Db 80 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCCCTAAGAAAGAGTTCTCGC 139
QY 121 CGGTTTGTCTATCTTCCAAATCCAGTACCTCTGATATTTGGAAAATGTATAAAGCAGCAG 180
Db 140 CGGTTTGTCTATCTTCCAAATCCAGTACCTCTGATATTTGGAAAATGTATAAAGCAGCAG 199
QY 181 GCTTCTCTTGTGACAGCAGAGAGGTCGACTTATCAAAAGGATCTCCCTCACTGGAACAG 240
Db 200 GCTTCTCTTGTGACAGCAGAGAGGTCGACTTATCAAAAGGATCTCCCTCACTGGAACAG 259
QY 241 CTTAAAGCAGATGAGAGTACTTCTCTCACATCTTAGCCCTTTTGTGAGCCAGTGAT 300
Db 260 CTTAAAGCAGATGAGAGTACTTCTCTCACATCTTAGCCCTTTTGTGAGCCAGTGAT 319
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGTCAGAGGCT 360
Db 320 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGTCAGAGGCT 379
QY 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db 380 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 439
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 440 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 499
QY 481 ACCATGCCCTATGTTAAGAAAGGAGCAGATTTGGGCTTGGATGGATAGCAGATAGAAA 540
Db 500 ACCATGCCCTATGTTAAGAAAGGAGCAGATTTGGGCTTGGATGGATAGCAGATAGAAA 559
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGTCTTCTTCTCAGGA 600
Db 560 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGTCTTCTTCTCAGGA 619
QY 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGCTTTATGCGCAGACTCACTTTTTC 660
Db 620 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGCTTTATGCGCAGACTCACTTTTTC 679
QY 661 AATGAACTCATCAGCAGATGAGAGCTTCACTGTGACTTTGCTTGGCTGATGTTCCAA 720

```
Db      |||||||
680 AATGAACTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCCTGATGTTCCAA 739
QY      |||||||
721 TACTTAGTAAATAAGCCTTCAGAGAAGAGGTCAGGAGATCATTTGTTGATGCTGTCTCAA 780
Db      |||||||
740 TACTTAGTAAATAAGCCTTCAGAGAAGAGGTCAGGAGATCATTTGTTGATGCTGTCTCAA 799
QY      |||||||
781 ATTGACGAGGAGTTTTTAACAGAAGCCTTGCCAGTTGGCCCTCATTTGGAATGAATTGCATT 840
Db      |||||||
800 ATTGACGAGGAGTTTTTAACAGAAGCCTTGCCAGTTGGCCCTCATTTGGAATGAATTGCATT 859
QY      |||||||
841 TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
Db      |||||||
860 TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 919
QY      |||||||
901 AAGGTTTTTCAGGCAGAAAATCCTTTTGATTTTATGMAAACATTTCTTTAGAGGAAAA 960
Db      |||||||
920 AAGGTTTTTCAGGCAGAAAATCCTTTTGATTTTATGMAAACATTTCTTTAGAGGAAAA 979
QY      |||||||
961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020
Db      |||||||
980 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1039
QY      |||||||
1021 ACAGATAACGTCTTTCACCTTGGATGCAGATTTT 1053
Db      |||||||
1040 ACAGATAACGTCTTTCACCTTGGATGCAGATTTT 1072
```

Search completed: October 30, 2005, 01:51:28
Job time : 3263 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2005, 21:38:30 ; Search time 440.5 Seconds
(without alignments)
14150.937 Million cell updates/sec

Title: US-10-698-228-2

Perfect score: 1053

Sequence: 1 atggcgaccggaagcc.....tcaccttgatgcagatttt 1053

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1053	100.0	1053	4	Aaf32438 Human rib
2	1053	100.0	2596	3	Aaa12411 CDNA enco
3	1053	100.0	4955	4	Aaf32439 Human rib
4	1053	100.0	4955	13	Adr24210 Breast ca
5	1051.4	99.8	1053	4	Aaf32447 Human rib
6	1051.4	99.8	1081	4	Aaf32440 Human rib
7	1049.8	99.7	1601	4	Aah14924 Human cDN
8	650.6	61.8	706	13	Adq57092 Novel can
9	578.4	54.9	1989	3	Aac78111 Human can
10	578.4	54.9	2215	12	Adk70302 Respirato
11	578.4	54.9	2216	10	Adj56536 Murine cD
12	578.4	54.9	2482	4	Aas44917 Human con
13	578.4	54.9	2500	6	Ab165414 Lung can
14	578.4	54.9	2500	6	Ab166517 Lung can
15	578.4	54.9	2500	6	Ab165859 Lung can
16	578.4	54.9	2500	8	Abx10335 DNA enco
17	578.4	54.9	2500	11	Adi32044 Human cDN
18	578.4	54.9	2500	12	Adn04443 Antipsori
19	578.4	54.9	2500	12	Adq09273 Human RRM
20	578.4	54.9	2500	13	Acn37637 Tumour-as

ALIGNMENTS

RESULT 1

AAF32438
ID AAF32438 standard; cDNA; 1053 BP.

XX AAF32438;

DT 18-APR-2001 (first entry)

DE Human ribonucleotide reductase TP53R2H nucleotide sequence SEQ ID NO:2.

Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.

OS Homo sapiens.

PN WO200100799-A1.

XX 04-JAN-2001.

XX 27-JUN-2000; 2000WO-JF004189.

PR 28-JUN-1999; 99JP-00181131.

PR 06-JUL-1999; 99JP-00192391.

PR 21-JAN-2000; 2000JP-00017770.

XX (TAKE) TAKEDA CHEM IND LTD.

XX (NAKA/) NAKAMURA Y.

XX Nakamura Y, Arakawa H, Tanaka H;

XX WPI; 2001-112446/12.

XX P-PSDB; AAB69050.

XX Ribonucleotide reductase involved in DNA repair and DNA encoding it, for diagnosis, treatment and prevention of cancer.

XX Claim 5; Fig 1-3; 102pp; Japanese.

XX The present sequence encodes a human ribonucleotide reductase designated TP53R2H. TP53R2H has cytostatic activity. TP53R2H is part of the DNA repair mechanism and its activity is induced by p53. It can be used for the treatment, prevention and diagnosis of a wide range of cancers

XX Sequence 1053 BP; 314 A; 192 C; 246 G; 301 T; 0 U; 0 Other;


```
Db 56 ATGGGCGACCCGGAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 115
QY 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 116 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 175
QY 121 CGGTTTGTCTATCTTTCAATCCAGTACCCCTGATATTTGGAAATGTATATAACAGGCACAG 180
Db 176 CGGTTTGTCTATCTTTCAATCCAGTACCCCTGATATTTGGAAATGTATATAACAGGCACAG 235
QY 181 GCTTCTCTTCTGGACAGCAGAGAGTGCATTTATCAAGAGATCTCCCTCACTGGAACAAG 240
Db 236 GCTTCTCTTCTGGACAGCAGAGAGTGCATTTATCAAGAGATCTCCCTCACTGGAACAAG 295
QY 241 CTTTAAAGCAGATCAGAGTACTTCTATCTCTCATCTTACCTTTAGCCCTTTTTCAGCCAGTAT 300
Db 296 CTTTAAAGCAGATCAGAGTACTTCTATCTCTCATCTTACCTTTAGCCCTTTTTCAGCCAGTAT 355
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCT 360
Db 356 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCT 415
QY 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAATTTTCACTCAGAGATGTACAGTTTG 420
Db 416 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAATTTTCACTCAGAGATGTACAGTTTG 475
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAATTTTATTTAATGCAATGAA 480
Db 476 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAATTTTATTTAATGCAATGAA 535
QY 481 ACCATCCCTCTATGTTAAGAAAAGCAGATTTGGCCCTTGGATGGATAGCAGATAGAA 540
Db 536 ACCATCCCTCTATGTTAAGAAAAGCAGATTTGGCCCTTGGATGGATAGCAGATAGAA 595
QY 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAAGAGTTTCTTCTCAGGA 600
Db 596 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAAGAGTTTCTTCTCAGGA 655
QY 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGCTTATGACAGAGACTCACTTTTCC 660
Db 656 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGCTTATGACAGAGACTCACTTTTCC 715
QY 661 AATGAATCATCAGCAGATGAGGACTTCACTGTGACTTTTCTTGCCTGATGTTCCAA 720
Db 716 AATGAATCATCAGCAGATGAGGACTTCACTGTGACTTTTCTTGCCTGATGTTCCAA 775
QY 721 TACTTGTAAATTAAGCTTTCAGAGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAA 780
Db 776 TACTTGTAAATTAAGCTTTCAGAGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAA 835
QY 781 ATTGACGAGAGTTTTTAAACAGAGCCTTTCAGGTTGGCTCATTTGGAATGCAAT 840
Db 836 ATTGACGAGAGTTTTTAAACAGAGCCTTTCAGGTTGGCTCATTTGGAATGCAAT 895
QY 841 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTAATTGTGGAACCTTGAATTTCA 900
Db 896 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTAATTGTGGAACCTTGAATTTCA 955
QY 901 AAGGTTTTTTCAGGACAGAAATCCTTTTGTATTTTATGGAACAACTTTCTTTAGAGGAAA 960
Db 956 AAGGTTTTTTCAGGACAGAAATCCTTTTGTATTTTATGGAACAACTTTCTTTAGAGGAAA 1015
QY 961 ACAAAATTTCTTTGAGAAACAGGTTTCAGAGTATCAGCGTTTTCAGGTTATGCGAGAAACC 1020
Db 1016 ACAAAATTTCTTTGAGAAACAGGTTTCAGAGTATCAGCGTTTTCAGGTTATGCGAGAAACC 1075
QY 1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053
Db 1076 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1108
```

RESULT 3

```
AAF32439
ID AAF32439 standard; cDNA; 4955 BP.
XX
AC AAF32439;
XX
DT 18-APR-2001 (first entry)
XX
DE Human ribonucleotide reductase related nucleotide sequence SEQ ID NO:3.
XX
KW Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.
XX
OS Homo sapiens.
XX
PN WO200100799-A1.
XX
PD 04-JAN-2001.
XX
PF 27-JUN-2000; 2000WO-JP004189.
XX
PR 28-JUN-1999; 99JP-00181131.
PR 06-JUL-1999; 99JP-00192391.
PR 21-JAN-2000; 2000JP-00017770.
XX
(TAKE ) TAKEDA CHEM IND LTD.
PA (NAKA/) NAKAMURA Y.
XX
PI Nakamura Y, Arakawa H, TanaKa H;
XX
WPI; 2001-112446/12.
XX
R Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
diagnosis, treatment and prevention of cancer.
XX
Example 2; Page 87-90; 102pp; Japanese.
XX
The present invention describes a human ribonucleotide reductase
designated TP53R2H. TP53R2H has cytostatic activity. TP53R2H is part of
the DNA repair mechanism and its activity is induced by p53. It can be
used for the treatment, prevention and diagnosis of a wide range of
cancers. The present sequence represents a human ribonucleotide reductase
related sequence which is used in an example from the present invention
XX
SQ Sequence 4955 BP; 1526 A; 803 C; 999 G; 1627 T; 0 U; 0 Other;
```

```
Query Match 100.0%; Score 1053; DB 4; Length 4955;
Best Local Similarity 100.0%; Pred. No. 2.1e-290;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGCGACCCGGAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 245 ATGGGCGACCCGGAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
QY 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 305 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 364
QY 121 CGGTTTGTCTATCTTTCAATCCAGTACCCCTGATATTTGGAAATGTATATAACAGGCACAG 180
Db 365 CGGTTTGTCTATCTTTCAATCCAGTACCCCTGATATTTGGAAATGTATATAACAGGCACAG 424
QY 181 GCTTCTCTTGGACAGCAGAGAGTGCATTTATCAAGGATCTCCCTCACTGGAACAAG 240
Db 425 GCTTCTCTTGGACAGCAGAGAGTGCATTTATCAAGGATCTCCCTCACTGGAACAAG 484
QY 241 CTTTAAAGCAGATCAGAGTACTTCTATCTCTCATCTTTAGCCCTTTTTCAGCCAGTAT 300
Db 485 CTTTAAAGCAGATCAGAGTACTTCTATCTCTCATCTTTAGCCCTTTTTCAGCCAGTAT 544
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCT 360
Db 545 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCT 604
QY 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAATTTTCACTCAGAGATGTACAGTTTG 420
```

Db 605 CGCTGTTTCTATGGCTTCAAAATCTCATCGAATGTTCACTCAGAGATGTACAGTTTG 664
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 480
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 724
Qy 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGCGATGGATAGCAGATAGAAA 540
Db 725 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGCGATGGATAGCAGATAGAAA 784
Qy 541 TCTACTTTTGGGAAAGAGTGGGCTTTGCTGTGTAGAGAGTTTCTTCTCAGGA 600
Db 785 TCTACTTTTGGGAAAGAGTGGGCTTTGCTGTGTAGAGAGTTTCTTCTCAGGA 844
Qy 601 TCTTTTGTCTATATCTGGCTTAAGAAAGAGAGTCTTATGCCAGGACTCATTCTTCC 660
Db 845 TCTTTTGTCTATATCTGGCTTAAGAAAGAGAGTCTTATGCCAGGACTCATTCTTCC 904
Qy 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGGCTGTGTTCCAA 720
Db 905 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGGCTGTGTTCCAA 964
Qy 721 TACTTAGTAAATAAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTGTGCTGTCAAA 780
Db 965 TACTTAGTAAATAAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTGTGCTGTCAAA 1024
Qy 781 ATTGACGAGAGTTTAAACAGAAAGCTTCCAGTGTGGCCTCATTTGGAATGAATTGATT 840
Db 1025 ATTGACGAGAGTTTAAACAGAAAGCTTCCAGTGTGGCCTCATTTGGAATGAATTGATT 1084
Qy 841 TTGATCAAAACAGTACATTTAGTTTGTAGCTGACAGATTAATCTGTGAACTTGGATTCTCA 900
Db 1085 TTGATCAAAACAGTACATTTAGTTTGTAGCTGACAGATTAATCTGTGAACTTGGATTCTCA 1144
Qy 901 AAGGTTTTTTCAGGAGAGAAATCCTTTTGTATTTTATGGAAACATTTCTTTAGAGAGAAA 960
Db 1145 AAGGTTTTTTCAGGAGAGAAATCCTTTTGTATTTTATGGAAACATTTCTTTAGAGAGAAA 1204
Qy 961 ACAATTTCTTTGAGAAACAGATTTTCAGATATCAGGTTTTCAGATTTATGGCAGAAACC 1020
Db 1205 ACAATTTCTTTGAGAAACAGATTTTCAGATATCAGGTTTTCAGATTTATGGCAGAAACC 1264
Qy 1021 ACAGATAAGCTCTTCACTTGGATGAGATTTT 1053
Db 1265 ACAGATAAGCTCTTCACTTGGATGAGATTTT 1297

RESULT 4

ADR24210

ID ADR24210 standard; DNA; 4955 BP.

XX

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

PI Van't Veer LJ, He Y;

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

XX WPI; 2004-593473/57.

Db 905 AATGAACATCATCAGCAGAGATGAAGGACCTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 964
Qy 721 TACTTAGTAATAAGCCTTCAGAGAAGAGGCTCAGGAGATCATTTGTTGATGCTGTCAAA 780
Db 965 TACTTAGTAATAAGCCTTCAGAGAAGAGGCTCAGGAGATCATTTGTTGATGCTGTCAAA 1024
Qy 781 ATTGAGCAGAGATTTTAAACAGAGCCTTGCAGTTGGCCTCATTTGGAATGGAATTGCATT 840
Db 1025 ATTGAGCAGAGATTTTAAACAGAGCCTTGCAGTTGGCCTCATTTGGAATGGAATTGCATT 1084
Qy 841 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTACTTGTGAACTTGGATTCTCA 900
Db 1085 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTACTTGTGAACTTGGATTCTCA 1144
Qy 901 AAGGTTTTTTCAGGCAGAGAAATCCCTTTGATTTTATGGAACCATTTCTTTAGAGGAAAA 960
Db 1145 AAGGTTTTTTCAGGCAGAGAAATCCCTTTGATTTTATGGAACCATTTCTTTAGAGGAAAA 1204
Qy 961 ACAAATTTCTTGGAGAAACAGGTTTTCAGAGTATCAGCGTTTTCAGGTTATGCGAGAAACC 1020
Db 1205 ACAAATTTCTTGGAGAAACAGGTTTTCAGAGTATCAGCGTTTTCAGGTTATGCGAGAAACC 1264
Qy 1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053
Db 1265 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1297

RESULT 5

AAF32447
ID AAF32447 standard; cDNA; 1053 BP.
AC AAF32447;
XX
XX
DT 18-APR-2001 (first entry)
DE Human ribonucleotide reductase related nucleotide sequence SEQ ID NO:12.
XX Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.
XX Homo sapiens.
XX
XX WO200100799-A1.
XX
XX PD 04-JAN-2001.
XX
XX PF 27-JUN-2000; 2000MO-JP004189.
XX
XX PR 28-JUN-1999; 99JP-00181131.
XX PR 06-JUL-1999; 99JP-00192391.
XX PR 21-JAN-2000; 2000JP-0001770.
XX
XX (TAKE) TAKEDA CHEM IND LTD.
XX (NAKA/) NAKAMURA Y.
XX
XX PA Nakamura Y, Arakawa H, Tanaka H;
XX
XX PI WPI; 2001-112446/12.
XX
XX DR Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
XX diagnosis, treatment and prevention of cancer.
XX
XX PS Claim 6; Page 96-97; 102pp; Japanese.
XX
XX CC The present invention describes a human ribonucleotide reductase
XX designated TP53R2H. TP53R2H has cytosolic activity. TP53R2H is part of
XX the DNA repair mechanism and its activity is induced by p53. It can be
XX used for the treatment, prevention and diagnosis of a wide range of
XX cancers. The present sequence represents a human ribonucleotide reductase
XX related sequence which is given in the exemplification of the present
XX invention
XX
XX Sequence 1053 BP; 314 A; 191 C; 246 G; 302 T; 0 U; 0 Other;

```

Db      1021 ACAGATAACGTCCTTCACCTTGGATGCAGATTTT 1053
|||||
RESULT 6
AAF32440
ID      AAF32440 standard; cDNA; 1081 BP.
XX
AC      AAF32440;
XX
DT      18-APR-2001 (first entry)
XX
DE      Human ribonucleotide reductase related nucleotide sequence SEQ ID NO:4.
XX
KW      Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.
XX
OS      Homo sapiens.
XX
PN      WO200100799-A1.
XX
PD      04-JAN-2001.
XX
PF      27-JUN-2000; 2000WO-JP004189.
XX
PR      28-JUN-1999; 99JP-00181131.
XX
PR      06-JUL-1999; 99JP-00192391.
XX
PR      21-JAN-2000; 2000JP-00017770.
XX
(PAKE ) TAKEDA CHEM IND LTD.
PA      (NAKA/) NAKAMURA Y.
XX
PI      Nakamura Y, Arakawa H, Tanaka H;
XX
XX      WPI; 2001-112446/12.
XX
DR      Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
PT      diagnosis, treatment and prevention of cancer.
XX
XX      Example 2; Page 91; 102pp; Japanese.
XX
CC      The present invention describes a human ribonucleotide reductase
CC      designated TP53R2H. TP53R2H has cytosolic activity. TP53R2H is part of
CC      the DNA repair mechanism and its activity is induced by p53. It can be
CC      used for the treatment, prevention and diagnosis of a wide range of
CC      cancers. The present sequence represents a human ribonucleotide reductase
CC      related sequence which is used in an example from the present invention
XX
SQ      Sequence 1081 BP; 322 A; 199 C; 252 G; 308 T; 0 U; 0 Other;

Query Match          99.8%; Score 1051.4; DB 4; Length 1081;
Best Local Similarity 99.9%; Pred. No. 2.9e-290;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1  ATGGCGACCCGGAAAGCGCGGAGCGCGGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db      |||||
Db      20  ATGGCGACCCGGAAAGCGCGGAGCGCGGGCTGGATCAGGATGAGAGATCATCTTCA 79
QY      61  GACACCAACGAAAGTGAATTAAGTCAAATGAAGAGCCACTCCCTAAGAAAGAGTTCTCGC 120
Db      |||||
Db      80  GACACCAACGAAAGTGAATTAAGTCAAATGAAGAGCCACTCCCTAAGAAAGAGTTCTCGC 139
QY      121 CGTTTGTGATCTTTCAATCCAGTACCTCGATATTGGGAAATGTATTAACAGGACAG 180
Db      |||||
Db      140 CGGTTTGTGATCTTTCCATCCAGTACCTCGATATTGGGAAATGTATTAACAGGACAG 199
QY      181 GCTTCTCTTGGACAGCAGAGAGGTGACTTCAAGGATCTCCCTCACTGGAACAAG 240
Db      |||||
Db      200 GCTTCTCTTGGACAGCAGAGAGGTGACTTCAAGGATCTCCCTCACTGGAACAAG 259
QY      241 CTTTAAAGCAGATGAGAAGTACTTCTCATCTCTCAATCTTAGCCCTTTTTCAGGCCAGTGAT 300
Db      |||||
Db      260 CTTTAAAGCAGATGAGAAGTACTTCTCATCTCTCAATCTTAGCCCTTTTTCAGGCCAGTGAT 319

```

```

QY      301 GGAATTTAAATGAAAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
Db      |||||
Db      320 GGAATTTGTAATGAAAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 379
QY      361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAATGTTTCACTCAGAGATGTACAGTTTG 420
Db      |||||
Db      380 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAATGTTTCACTCAGAGATGTACAGTTTG 439
QY      421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTAAATGCAATTGAA 480
Db      |||||
Db      440 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTAAATGCAATTGAA 499
QY      481 ACCATGCCCTATGTTTAAAGAAAAAGCAGATTGGCCCTTCGATGATGATGATGATGATGAT 540
Db      |||||
Db      500 ACCATGCCCTATGTTTAAAGAAAAAGCAGATTGGCCCTTCGATGATGATGATGATGATGAT 559
QY      541 TCTACTTTTGGGGAAGAGTGGTGGCTTTCCTGCTGTAGAAAGAGTTTCTTCTTCAGGA 600
Db      |||||
Db      560 TCTACTTTTGGGGAAGAGTGGTGGCTTTCCTGCTGTAGAAAGAGTTTCTTCTTCAGGA 619
QY      601 TCTTTTGTCTATATTCTGCTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
Db      |||||
Db      620 TCTTTTGTCTATATTCTGCTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTTC 679
QY      661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTCTGCTGATGTTTCCAA 720
Db      |||||
Db      680 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTCTGCTGATGTTTCCAA 739
QY      721 TACTTAGTAATAAGCCTTCAGAAAGAGGTCAAGGAGATCATTTGTTGATGCTGTGCAAA 780
Db      |||||
Db      740 TACTTAGTAATAAGCCTTCAGAAAGAGGTCAAGGAGATCATTTGTTGATGCTGTGCAAA 799
QY      781 ATTGAGCAGGAGTTTAAACAGAGAGCCTTGGCCAGTTGGCCCTCATTTGGAATGAATTGCATT 840
Db      |||||
Db      800 ATTGAGCAGGAGTTTAAACAGAGAGCCTTGGCCAGTTGGCCCTCATTTGGAATGAATTGCATT 859
QY      841 TTGATGAAACAGTACATTTAGTTTGTAGCTCAGCAGATTTACTTTGTGAACTTGGATTCTCA 900
Db      |||||
Db      860 TTGATGAAACAGTACATTTAGTTTGTAGCTCAGCAGATTTACTTTGTGAACTTGGATTCTCA 919
QY      901 AAGGTTTTTTCAGGCGAGAAAAATCTTTTGAATTTTATGGAATAACATTTCTTTAGAGGAAAA 960
Db      |||||
Db      920 AAGGTTTTTTCAGGCGAGAAAAATCTTTTGAATTTTATGGAATAACATTTCTTTAGAGGAAAA 979
QY      961 ACAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGGTTATGCGCAGAAACC 1020
Db      |||||
Db      980 ACAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGGTTATGCGCAGAAACC 1039
QY      1021 ACAGATAACGTCCTTCACCTTGGATGCAGATTTT 1053
Db      |||||
Db      1040 ACAGATAACGTCCTTCACCTTGGATGCAGATTTT 1072

RESULT 7
AAH14924
ID      AAH14924 standard; cDNA; 1601 BP.
XX
AC      AAH14924;
XX
DT      26-JUN-2001 (first entry)
XX
DE      Human cDNA sequence SEQ ID NO:12810.
XX
KW      Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS      Homo sapiens.
XX
PN      EP1074617-A2.
XX
PD      07-FEB-2001.
XX
PF      28-JUL-2000; 2000EP-00116126.
XX

```


XX (GENE-) GENE LOGIC INC.
 PA (PFIZ) PFIZER PROD INC.
 XX
 PI Diggins JC, Porter M, Wei T;
 XX WPI; 2004-561890/54.
 DR
 XX
 XX New isolated nucleic acid molecule, useful for drug screening and
 PT toxicity assays or for assessing the impact, including toxicity, of a
 PT compound, pharmaceutical agent or environmental pollutant on a cell or
 PT living organism.
 XX
 PS Claim 1; SEQ ID NO 8394; 4lpp; English.
 XX
 CC This invention is related to a novel isolated canine nucleic acid
 CC sequences and the construction of canine microarrays containing a
 CC significant portion of the canine genome. The isolated canine nucleic
 CC acid sequences of the invention may be useful for drug screening and
 CC toxicity assays. The invention is therefore useful for assessing the
 CC impact, including toxicity, of a compound, pharmaceutical agent or
 CC environmental pollutant on a cell or living organism. The methods are
 CC useful for detecting genes that are up- or down-regulated in canines in a
 CC disease state. The sequences are useful as diagnostic agents or markers
 CC to detect a cellular response in a sample individually or as part of a
 CC gene expression profile. It is also useful as a target for agents that
 CC modulate gene expression or activity. The database is useful for
 CC producing electronic Northernblots that allow the user to determine the cell
 CC type or tissue in which a given gene is expressed and to allow
 CC determination of the abundance or expression level of a given gene in a
 CC particular tissue or cell. The methods are useful for determining the
 CC similarity of a toxic response to one or more individual compounds. The
 CC methods are useful for predicting at least one toxic response or the
 CC likelihood that a compound or test agent will induce various specific
 CC pathologies such as those of the liver (liver necrosis, fatty liver
 CC disease, protein adduct formation or hepatitis), those of the kidney,
 CC heart, brain or testes, or other pathologies associated with at least one
 CC of the toxins. The methods are also useful for predicting or elucidating
 CC the potential cellular pathways influenced, induced or modulated by the
 CC compound or test agent due to the similarity of the expression profile
 CC compared to the profile induced by a known toxin. The present sequence is
 CC that of a canine DNA sequence which was claimed for use during the
 CC production of a canine microarray of the invention.
 XX
 SQ Sequence 706 BP; 208 A; 113 C; 158 G; 227 T; 0 U; 0 Other;
 Query Match 61.8%; Score 650.6; DB 13; Length 706;
 Best Local Similarity 95.2%; Pred. No. 1.4e-175;
 Matches 671; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
 349 GTTCCAGAGCTCGCTGTTTCTATGGCTTTTCAAATTTCTATCGAGAAATGTTCTACACGAG 408
 1 GTTCCAGAGCTCGCTGTTTCTATGGCTTTTCAAATTTCTATCGAGAAATGTTCTACACGAG 60
 409 ATCTACAGTTTGTCTAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTT 468
 61 ATGTACAGTTTGTCTAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTT 120
 469 AATGCAATTAACCACTGCTATGTTAAGAAAGGAGAGATGCGGCTTGGCATGGATA 528
 121 AATGCAATTAACCACTGCTATGTTAAGAAAGGAGAGATGCGGCTTGGCATGGATA 180
 529 GCAGATAGAAATCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAGAGATT 588
 181 GAAGATAGAAATCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAGAGATT 240
 589 TTCTTCTCAGATCTTTGCTGCTATATTTCTGCTAAGAGAGAGATCTTATGCCAGGA 648
 241 TTCTTCTCAGATCTTTGCTGCTATATTTCTGCTAAGAGAGAGATCTTATGCCAGGA 300
 649 CTCACTTTTCCAAATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGC 708
 301 CTCACTTTTCCAAATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGC 360

709 CTGATGTTCCAAATCTAGTAAATAAGCCTTTCAGAGAAAGGGTCAGGAGATCATTTGTT 768
 361 CTGATGTTCCAAATCTAGTAAATAAGCCTTTCAGAGAAAGGGTCAGGAGATCATTTGTT 420
 769 GATGCTGTCAAAATTTGAGCAGGAGTTTTAAACAGAAAGCCTTGCCAGTTGGCCTCATTTGA 828
 421 AATGCTGTGAAATTTGAGCAGGAGTTTTAAACAGAAAGCCTTGCCAGTTGGCCTCATTTGA 480
 829 ATCAATTTGCAATTTTGTATGAAACAGTACATTTAGTGTTCAGCTGACAGATTAATTTGTGAA 888
 481 ATCAATTTGCTGTTTGTATGAAACAGTATATTTAGTGTTCAGCTGACAGATTAATTTGTGAA 540
 889 CTTGGATTTCTCAAAAGGTTTTCAGGAGAAATCCTTTTGATTTTATGGAACCAATTTCT 948
 541 CTTGGATTTCTCAAAAGGTTTTCAGGAGAAATCCTTTTGATTTTATGGAACCAATTTCT 600
 949 TTAGAAGGAAACCAAAATTTCTTTTGAAGAAAGAGTTTCAGAGTATCAGCGTTTTCAGTT 1008
 601 TTAGAAGGAAACCAAAATTTCTTTTGAAGAAAGAGTTTCAGAGTATCAGCGTTTTCAGTT 660
 1009 ATGCGAGAAACCAAGATACGTTCTTCACTTGGATGCGAGATTTT 1053
 661 ATGCGAGAAACCAAGATACGTTCTTCACTTGGATGCGAGATTTT 705

RESULT 9
 AAC78111
 ID AAC78111 standard; cDNA; 1989 BP.
 XX
 AC AAC78111;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human cancer associated gene sequence SEQ ID NO:505.
 XX
 KW Human; cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cytostatic; proliferative; vulnery; immunomodulator;
 KW antidiabetic; antisthmatic; antirheumatic; antiarthritic; antiviral;
 KW antiinflammatory; anticholel; antiallergic; antibacterial; cardiant;
 KW dermatologic; neuroprotective; thrombolytic; coagulant; nootropic;
 KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
 KW allergic reaction; graft versus host disease; organ rejection;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;
 KW neurological disease; drug screening; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200055350-A1.
 XX
 PD 21-SEP-2000.
 XX
 PF 08-MAR-2000; 2000WO-US005882.
 XX
 PR 12-MAR-1999; 99US-0124270P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX
 XX WPI; 2000-587533/55.
 DR P-PSDB; AAB43902.
 XX
 PT Novel isolated nucleic acids comprising sequences encoding peptides
 useful for treating or diagnosing e.g. cancer.
 PS
 Claim 1; Page 1035-1036; 2352pp; English.
 XX
 CC AAB7607 to AAC78448 encode the human cancer associated proteins given in
 CC AAB4398 to AAB44239. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities
 CC include: cytostatic; proliferative; vulnery; immunomodulator;

CC anti-diabetic; antiasthmatic; antirheumatic; antiarthritic;
 CC anti-inflammatory; antithyroid; anti-allergic; antibacterial; antiviral;
 CC dermatological; neuroprotective; cardiac; thrombolytic; coagulant;
 CC neotrophic; vasotropic; antiposrotic and antiangiogenic. The
 CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of
 CC the present invention
 XX
 SQ Sequence 1989 BP; 525 A; 439 C; 476 G; 544 T; 0 U; 5 Other;

Query Match 54.9%; Score 578.4; DB 3; Length 1989;

Best Local Similarity 75.0%; Pred. No. 1.1e-154;

Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

QY 89 ATGAAGAGCCACTCTTAAGAAAGAGTTCTCGCGGTTTGTCTATCTTTTCCAATCCAGTACC 148
 DB |||||
 QY 149 CTGATATTGGAAATGTATAACAGCAGCAGGCTTCTCTCGACAGCAGAGAGGTG 208
 DB |||||
 QY 340 ATGATATCTGGCAGATGTATAAGAGCAGAGGCTTCTTTTGGACCGCGGAGGAGTGG 399
 DB |||||
 QY 209 ACTTATCAAGAGTCTCCCTCACTGGAACAGCTTAAGCAGATGAGAATCTTCTATCT 268
 DB |||||
 QY 400 ACCTCTCCAGGACATTCAGCATCTGGGAATCCCTGAAACCCGAGGAGAGATATTTATAT 459
 DB |||||
 QY 269 CTCACATCTTAGCTTTTTCAGCCAGATGGAATTTGAAATGAAATTTGGTGGAGC 328
 DB |||||
 QY 460 CCCATGTTCTGGCTTCTTTTGCAGCAGCGATGGCATAGTAATGAACTTGGTGGAGC 519
 DB |||||
 QY 329 GCTTTAGTCAGGAGGTGCGAGGTTCAGAGGCTCGCTGTTTCTATGCTTTCAAATTTCTCA 388
 DB |||||
 QY 520 GATTTAGCCAAAGAGTTCAGATTACAGAGCCGCTGTTTCTATGCTTCCAAATTTGCCA 579
 DB |||||
 QY 389 TCGAGATCTTCACTCAGAGATGTACAGTTTGTCTGTATAGACATTCATCAGATCCCA 448
 DB |||||
 QY 580 TGAAGAACATACATCTGAAATGTATAGTCTTCTTATGACACTTACATAAAGATCCCA 639
 DB |||||
 QY 449 AGAAAGGGAAATTTTATTATTAATGCAATTTGAACCATGCCATATGTTAAGAAAAAGCAG 508
 DB |||||
 QY 640 AGAAAGGGAAATTTCTTCAATGCCATTTGAACCATGCTTGTGTCAGAGAGAGCAG 699
 DB |||||
 QY 509 ATTGGGCTTGGATGGATGACAGATAGAAAATCTACTTTTGGGGAAGAGTGGTGGCCT 568
 DB |||||
 QY 700 ACTGGGCTTGGCTGGATGGGACAAAGAGGCTACCTATGCTGAACGTTGTAGCCT 759
 DB |||||
 QY 569 TTGCTGCTGTAGAGAGTTTCTTCTCAGATCTTTTGGTGTCTATATCTGGCTAAGA 628
 DB |||||
 QY 760 TTGCTGCTGTAGAGAGTTTCTTCTCAGATCTTTTGGTGTCTATATCTGGCTAAGA 819
 DB |||||
 QY 629 AGAGAGGCTTATGCCAGGACTCACATTTTCCAATGAATCTATCAGCAGAGATGAGGAC 698
 DB |||||
 QY 820 AAGAGGACTGATGCTGGCTGCTCACATTTCTTAATGAATCTTATAGCAGATGAGGGTT 879
 DB |||||
 QY 689 TTTCACTGTGACTTTGCTTGCCTGATGTTCCAAATCTAGTAATAAGCCCTTCAGAGAAA 748
 DB |||||
 QY 880 TACACTGTGATTTGCTTGCCTGATGTTCAACACCTGCTGATACACCAACCTGAGGAGA 939
 DB |||||
 QY 749 GGGTCAGGAGATCATTTGTGATGCTGTCAAAATTCAGCAGGAGTTTAAACAGAGCCT 808
 DB |||||
 QY 940 CAGTAAGAGAAATATATCAATGCTGTTTCGGATAGAACAGGAGTTTCTCACTGAGGCT 999
 DB |||||
 QY 809 TGCCAGTTGGCCTCATTTGGAATGAATTTGATGAAACAGTACATTTGATTTGTAG 868
 DB |||||

DB 1000 TGCCTGTGAAGCTCATTTGGGATGAATTCACCTCTAATGAAGCAATACATTGAGTTGTGG 1059
 QY 869 CTGACAGATTACTTGTGGAACCTTGATCTCAAGGTTTTTTCAGGCAGAAATCCTTTTG 928
 DB 1060 CAGACAGACTTATGCTGGAACCTGGGTTTTAGCAAGTTTTTCAGAGTAGAGAACCCATTG 1119
 QY 929 ATTTATGGAACAAATTTCTTTAGAAAGAAAAACAAATTTCTTTGAGAAACGAGTTTCAG 988
 DB 1120 ACTTTATGGAGATATTTCACTGGAGGAGAAAGACTTAACCTTTTGAAGAGAGTAGGCG 1179
 QY 989 AGTATCAGCGTTTTCAGTTATGCGAGAAACACAGATAACCTTTTACCTTTGGATGCGAG 1048
 DB 1180 AGTATCAGAGATGGAGTGTATGCAAGTCCAAACAGAGAAATTTTACCTTTGGATGCTG 1239
 QY 1049 ATTT 1052
 DB 1240 ACTT 1243

RESULT 10
 ADK70302

ID ADK70302 standard; cDNA; 2215 BP.

XX AC ADK70302;

XX DT 06-MAY-2004 (first entry)

XX DE Respiratory disease differentially expressed cDNA #38.

XX KW ds; gene; cytostatic; respiratory; antiasthmatic; gene therapy;
 KW differential gene expression; respiratory disorder; lung cancer;
 KW chronic obstructive pulmonary disease; emphysema; asthma.

XX OS Homo sapiens.

XX PN W02003101283-A2.

XX PD 11-DEC-2003.

XX PF 02-JUN-2003; 2003WO-US017409.

XX PR 04-JUN-2002; 2002US-0386005P.

XX PA (INCY-) INCYTE CORP.

XX PI Rickert PK, Krasnow R;

XX PS WPI; 2004-042945/04.

XX PT New combination comprising cDNAs and proteins that are differentially
 PT expressed in respiratory disorders, useful for diagnosing or treating
 PT respiratory diseases e.g. lung cancer, chronic obstructive pulmonary
 PT diseases or asthma.

PS Claim 1; SEQ ID NO 38; 343pp; English.

XX The invention relates to cDNA sequences that are differentially expressed
 CC in respiratory disorders or their complements or encoded proteins. The
 CC cDNAs and proteins are useful for diagnosing, treating or monitoring
 CC treatment of a subject with a respiratory disease including lung cancer,
 CC chronic obstructive pulmonary diseases, emphysema or asthma. The protein
 CC is also useful for screening molecules or compounds to identify at least
 CC one ligand which specifically binds the protein. It is also useful for
 CC preparing and purifying a polyclonal or monoclonal antibody. This
 CC sequence corresponds to a cDNA of the invention.

XX SQ Sequence 2215 BP; 593 A; 476 C; 523 G; 623 T; 0 U; 0 Other;

Query Match 54.9%; Score 578.4; DB 12; Length 2215;

Best Local Similarity 75.0%; Pred. No. 1.2e-154;

Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

Qy	89	ATGAAGGCCACTCTCTTAAGAAAGAGTTCTCGCCGGTTTGTGTCATCTTTCCAAATCCAGTACC	148
Db	283	AGGATGAGCCGCTGCTGAGAGAAACCCCGCGCTTTGTGTCATCTTTCCCAATCGAGTACC	342
Qy	149	CTGATATTTTCGAAAACTATAAACAGCAGCAGAGCTTCTCTTCGACAGCAGAGAAGAGTCG	208
Db	343	ATGATATCTGGCAGATGATTAAGAGGCAGAGGCTTCTTTTGGACCGCCGAGAGAGTGG	402
Qy	209	ACTTATCAAAAGGATCTCCCTCACTGGAAACAAGCTTTAAAGCAGATGAGAAGTACTTCATCT	268
Db	403	ACCTCTCCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT	462
Qy	269	CTCACATCTTAGCGTTTTCGACGCCAGTGATGGAAATGTAATGAAAAATTTGGTGGAGC	328
Db	463	CCCATGTTCTGGGCTTCTTTTCGACGAAGCATGGCATAGTAATGAAAACTTCGTGGAGC	522
Qy	329	GCTTTAGTCAGGAGTGACAGGTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAAATCTCA	388
Db	523	GATTTAGCCAAAGTTTCAGATTACAGAACCCGCTGTTTCTATGGCTTCCAAATTCGCCA	582
Qy	389	TCGAGAAATGTTCACTCAGAGATGATACAGTTTGCTGATAGACACTTATCATCAGAGATCCCA	448
Db	583	TGGAAAAACATACATTTCTGAAATGTATAGTCTTCTTATTCGACACTTACATAAAAAGATCCCA	642
Qy	449	AGAAAGGGAAATTTTATATTTAATGCAATTCGAACCAATGCCCTATGTTTAAAGAAAAAGCAG	508
Db	643	AAGAAAGGGAAATTTCTCTTCAATGCCAATGAAACGATGCCCTGTGTCAAGAGAAGGCAG	702
Qy	509	ATTGGGCTTTGCCATGGATAGCAGATAGAAAAATCTACTTTTGGGGAAAGAGTGGTGGCCT	568
Db	703	ACTGGGCTTTGCGCTGGATTTGGGGACAAAGAGGCTACTATGTTGAACGTGTTGTAGCCT	762
Qy	569	TTGCTGCTGTAGAAGAGAGTTTCTTCTCAGGATCTTTTGTGCTGTATATTTCTGGCTAAAGA	628
Db	763	TTGCTGCAGTGAAGGCAATTTCTTTTCCGGTCTCTTTTGGCTGCATATTTCTGGCTCAAGA	822
Qy	629	AGAGAGTCTTATGSCCAGGACTCACTTTTCCAAATGAACCTCATCAGCAGAGATGAAGGAC	688
Db	823	AACGAGACTGATGCTCGGCTCACATTTTCTAATGAACCTTATTAGCAGAGATGAGGGTT	882
Qy	689	TTCACTGTGACTTTGCTTGCTGATGTTCCAATACTTAGTAAATAAGCCTTCAGAAGAAA	748
Db	883	TACACTGTGATTTTGTCTGCTGATGTTCAACACACCTGGTACACAAACCATCGAGGAGA	942
Qy	749	GGGTCCGGAGATCATTTGTTGATGCTGTCAAAATTTAGCAGGAGTGTTTTAAACAGAAGCT	808
Db	943	GAGTAAGAGAAATAATTTATCAATGCTGTTCGGATAGAACAGGAGTTCCTCAGTGAGGCT	1002
Qy	809	TGCCAGTTGGCTCATTTGGAAATGAATTCATTTTGTAGTAAACAGTACATTTGAGTTTGTAG	868
Db	1003	TGCTGTGAAGCTCATTTGGGATGNAATTTGCACTTAATGAAGCAATACATTTGAGTTTGTG	1062
Qy	869	CTGACAGATTACTTGTGGAACCTTGATTCCTCAAGGTTTTCAGGCCAGAAAAATCCTTTTG	928
Db	1063	CAGACACATTAATGCTGGAACCTGGTTTATAGCAAGTTCCTCAGAGTAGAGAACCCATTG	1122
Qy	929	ATTTTATGGAACCATTTCTTTAGAAAGGAAAAAATAATTTCTTTTGAGAAACGAGTTTCAG	988
Db	1123	ACTTTATGGGAATATTTTCACTGGAAAGAAAGACTTAATCTTTTGGAAAGAGAGTAGCGC	1182
Qy	989	AGTATCAGCGTTTTCGAGTTTATGGCAGAAAAACCAAGATAAAGCTTCTCACCTTGGATCGAG	1048
Db	1183	AGTATCAGAGGATGGGAGTGATGCTCAAGTCCAAACAGAGAAATCTTTTACCTTGGATGCTG	1242
Qy	1049	ATTT 1052	
Db	1243	ACTT 1246	

AC	ADJ56536;
XX	
XX	06-MAY-2004 (first entry)
XX	
DE	Murine cDNA differentially expressed in MYCN activated cells SeqID 342.
XX	
KW	mouse; murine; differential expression; transactivator; proto-oncogene;
KW	neuroblastoma; small cell lung cancer; cystostatic; gene therapy; ss;
KW	MYCN activated cell.
XX	
OS	Mus musculus.
XX	
XX	US20031119009-A1.
PN	
XX	
XX	26-JUN-2003.
PD	
XX	
PF	25-FEB-2002; 2002US-00084817.
XX	
XX	23-FEB-2001; 2001US-0270784P.
PR	
XX	(STUA/) STUART S G.
PA	(NUCH/) NUCHTERN J G.
PA	(PLON/) PLON S E.
PA	(SHOH/) SHOHEIT J M.
XX	
XX	Stuart SG, Nuchtern JG, Plon SE, Shohet JM;
PI	
XX	WPI, 2003-635698/60.
XX	
DR	
XX	
PT	New genes regulated by MYCN activation, useful in gene therapy.
PT	particularly for treating a subject with e.g. neuroblastoma or other
PT	cancers, or for diagnosing, staging or monitoring the treatment of the
PT	cancer.
XX	
PS	Claim 1; SEQ ID NO 342; 27pp; English.
XX	
XX	This invention relates to novel isolated cDNAs that are differentially
CC	expressed in MYCN activated cells. Specifically, it refers to
CC	polynucleotide sequences that exhibit differential expression patterns in
CC	cells activated by the transactivator MYCN, where MYCN is a proto-
CC	oncogene that is amplified in neuroblastoma cells and is common in small
CC	cell lung cancers. The present invention describes these cDNA molecules
CC	as useful for in hybridisation assays to detect expression of nucleic
CC	acids (or complementary nucleic acids) in a present in a given sample, as
CC	well as for screening assays by identifying molecules or compounds that
CC	specifically bind the cDNA as a ligand and modulate function or activity.
CC	Accordingly, these compositions exhibit cytostatic activity and can also
CC	be used for gene therapy purposes. This polynucleotide sequence is a cDNA
CC	that is differentially expressed in MYCN activated cells, given in an
CC	exemplification of the invention. NOTE: This sequence does not appear in
CC	the printed specification but has been obtained in electronic format from
CC	the US Patent Office at
CC	ftp.seqdata.uspto.gov/sequence.html?DocID=20030119009.
XX	
SO	Sequence 2216 BP: 593 A: 476 C: 524 G: 623 T: 0 U: 0 Other:

RESULT 11
ADJ56536
ID ADJ5
XX

464	Db		CCCATGTTCTGGGCTTCTTTTCAGCAGCGATGGCATAGTAATGAACAACTTGGTGGAGC	523
329	Qy		GCTTTAGTCAGAGGTCAGGTTCCAGAGGTCGCTGTTCTATGGCTTTCAAATTTCTCA	388
524	Db		GATTTAGCCAAGATTCAGATTACAGAAAGCCGCTGTTCTATGGCTTCCAAATTCGCA	583
389	Qy		TCGAGAATGTTCTACTCAGAGATGTACAGTTTGCTGATAGACACTTACATCAGAGATCCCA	448
584	Db		TGGAAACATACATTCCTGAATGTATAGTCTCTTATTGTGACACTTACATAAAGATCCCA	643
449	Qy		AGAAAAGGGAATTTTATTATTCGAATGTGAAACCATGCCCTATGTTAAGAAAAAGCAG	508
644	Db		AAGAAAAGGGAATTTCTCTTCAATGCGATTGAAACGATGCGCTTGTTCTCAAGAAGAGGCG	703
509	Qy		ATTGGGCTTTCGATGGATACGATAGAAATCTACTTTTGGGGAAGAGTGTGGCCT	568
704	Db		ACTGGGCTTTCGCTGGATGGGACAAAGAGGTACCTATGSGTGAACGTGTTGAGCCT	763
569	Qy		TTGCTGCTGTGAAGAGGAGTTTCTTCTCAGGATCTTTTGCCTATATTTCTGGCTAAAGA	628
764	Db		TTGCTGCAGTGAAGGCATTTCTTTCCGGTCTTTTGCCTCGATATTTCTGGCTCAGA	823
629	Qy		AGAGAGTCTTATGCGCAGGACTCATCTTTTCCAAATGAACCTCATCAGCAGAGATCAAGAC	688
824	Db		AACGAGACTGATGCCCTGGCCTCACATTTCTTAATGAACCTATTAGCAGAGATGAGGTT	883
689	Qy		TTCACTGTGACTTTCCTTGCTGCTGATGTTCCAAATCTTAGTAAATAAGCCTTCAGAAGAA	748
884	Db		TACACTGTGATTTTTCCTTGGCTGATGTTCAAACACCTGGTACAAACCATCGGAGGAGA	943
749	Qy		GGGTCAAGGAGATCATTTGTTGCATCTGTCAAAATTTGACGAGGAGTTTTTAACAGAGCCT	808
944	Db		GAGTAAGAGAAATTAATTATCAATGCTGTTCCGATAGAACAGGAGTTCTCTACTGAGGCT	1003
809	Qy		TGCCAGTTGGCCTCATTTGGAAATGAATTTGCATTTTTCATGAAACAGTACATTCAGTTGTAG	868
1004	Db		TGCCTGTGAAGCTCATTTGGGATGAATTCGACTCTCTAATGAAGCAATACATTCAGTTGTGG	1063
869	Qy		CTGACAGATTACTTGTGGAACTTGGATTCTCAAAAGTTTTTCAGGCGAGAAATCCCTTTTG	928
1064	Db		CAGACAGACTTATGCTCGAACTGGGTTTTTACCAAGGTTTTTCAGAGTAGAAGAACCCATTTG	1123
929	Qy		ATTTTATGGAACAATTTCTTTAGAAGSAAAAACAATTTCTTTTGAGAAACGAGTTTCAG	988
1124	Db		ACTTTATGGAGAAATTTTCACTGGAAGGAAGAGCTAACTCTTTTGAAGAGAGATGAGCG	1183
989	Qy		AGTATCAGCGTTTTGTCAGTTTATGCGCAAGAAACACAGATTAACGCTCTTCACCTTTGGATGCG	1048
1184	Db		AGTATCAGAGGATGGGAGTGATGTCAGTCCAAACAGAGAAATCTTTTACCTTTGGATGCTG	1243
1049	Qy		ATTT 1052	
1244	Db		ACTT 1247	

RESULT 12

RESOL 12
AAS44917/C

ID AAS44917 standard; DNA: 2482 BP.
 AAS44917C

XX

AC AAS44917;

XX
XX
-----)

DT 18-DEC-2001 (first entry)

XX

DE Human con

XX

KW Mammal; h

KW mouse; Ch

KW cancer; 1:

KW nervous S:

KW angiogenic

genetic d

antibacterial; immunosuppressive; vasotropic; antiparkinsonian; neuroprotective; osteopathic; antidiabetic; antirheumatic; antiallergic; immunostimulant; analgesic; gene therapy.

Homo sapiens.
Synthetic.

WO200164834-A2.

07-SEP-2001.

26-FEB-2001: 2001WO-US004926.

28-FEB-2000: 2000IS-00515126.

28-FEB-2000; 2000US-00515126;
18-MAY-2000: 2000US-00577409.

18-MAY-2000; 2000US-00577409;
17-JUN-2000; 2000US-00597707;

17-JUN-2000; 2000US-0037707;
14-JUL-2000; 2000US-00616807;

14-SEP-2000; 2000US-00616807;
19-SEP-2000; 2000US-00664641.

(HYSE-) HYSEO INC.

Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;
Drimnac R;

WPI: 2001-589862/66.

WEI, ZOU-363802/
P-PSDB; AAU28017.

Novel polypeptides and nucleic acids obtained from cDNA libraries prepared from various human tissues, for diagnosis, treatment of cancer, neurological, inflammatory disorders and for use in arrays for detection

Claim 1; SEQ ID NO 514; 153pp; English.

Sequences AAS44576-AAS44919 represent full-length polynucleotides and contig polynucleotides encoding polypeptides of the invention. The DNA and protein sequences are useful for the treatment, diagnosis and prevention of various types of disorder in a mammalian subject such as a human, dog, monkey, mouse, hamster or rat. The disorders include cancers such as leukaemia, lymphoma and neuroblastoma, autoimmune disorders such as multiple sclerosis, connective tissue disease, rheumatoid arthritis, diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system disorders such as Parkinson's disease, Alzheimer's disease, Huntington's chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and Wernerke disease, inflammatory disorders such as nephritis, Crohn's disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory bowel disease. The sequences exhibit activity relating to angiogenesis, cell proliferation, cell differentiation, stem cell growth factor, activin or inhibin. Therefore, they can be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, accidental damage or genetic disorders. The sequences may also be used for regeneration of bone, cartilage, tendons and ligaments and in tissue repair and burn healing. Note: Some sequences for this patent did not form part of the printed specification, but were obtained in electronic format directly from WIPO at www.wipo.int/pub/published/pct sequences

Sequence 2482 BP; 707 A; 598 C; 522 G; 655 T; 0 U; 0 Other;

Query Match 54.9%: Score 578.4: DB 4: Length 2482:

Query Match 34.5%; score 378.4; DD 4.1
Best Local Similarity 75.0%; pred. No. 1.2e-154;

Best local similarity 75.0%, Recd: NO. 1.28-1.51,
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

QY 89 ATGAAGAGCCACTCCTAAGAAAGAGTTCTCGCCGGTTTGTCATCTTTCCAATCCAGTACC 148

0
1
2
3
4
5
6
7
8
9
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z

Db 2086 AGGATGAGCCGCTGCTGAGAGAAACCCCGCCGCTTTGTCACTTCCCCATCGAGTACC 2027

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

QY 149 CTGATATTTGGAAATGTATAAACAGGCACAGGCTTCCTTCTGGACAGCAGAGAGGTCG 208

Db 1966 ACCTCTCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT 1907
Qy 269 CTACATCTTTAGCCTTTTTCGAGCCAGTATGGAATTTGTAATGAAATTTGGTGGAGC 328
Db 1906 CCATGTTCTGGCTTTCTTTGAGCAAGCGATGGCATAGTAATAAGAACTTTGGTGGAGC 1847
Qy 329 GCTTTAGTCAGGAGGTCAGGCTCCAGAGCGCTGGCTTTCTATGGCTTTCAAAATTTCTCA 388
Db 1846 GATTTAGCAAGAAAGTTCAGATACAGAACCCGCTGTTCTATGGCTTCCAAATTTGCCA 1787
Qy 389 TCAGAAATGTTACTACAGAGATGACAGTTTGTGTATAGACACTTACATCAGAGATCCCA 448
Db 1786 TGGAAACATACATCTCGAAATGTATAGTCTTCTTATGACACTTACATAAAGATCCCA 1727
Qy 449 AGAAAGGGAATTTTATTTAATGCAATTCGAACCATGACCTATGTTTAAAGAAAAGCAG 508
Db 1726 AAGAAAGGGAATTTCTCTTCAATGCCATTCGAACGATGCTTGTGTCAAGAAGAGCAG 1667
Qy 509 ATTGGGCTTTCGATGATAGCAGATAGAAAATCTACTTTTGGGGAAGAGTGGTGGCCT 568
Db 1666 ACTGGGCTTTCGCTGATTTGGGACAAAGAGGCTACTATGTGAACGTGTTGTAGCCT 1607
Qy 569 TTGCTGCTGAGAGGAGTTTCTTCTCAGAGTCTTTTGTGCTATATTCGTGCTAAAGA 628
Db 1606 TTGCTGAGTGAAGGCAATTTCTTTTCCGGTCTTCTTGGCTCGATATCTGGCTCAAGA 1547
Qy 629 AGAGAGTCTTATGCCAGGACTCACTTTTCCATGAACTCATCAGCAGAGATGAAGGAC 688
Db 1546 AACGAGGACTGATGCTGGCTGCATTTCTAATGAACTTATAGCAGAGATGAGGGTT 1487
Qy 689 TTCACTGTGACTTGTGCTGATGTTCCAACTACTTGTAAATAAGCCCTTCAGAGAAA 748
Db 1486 TACACTGTGATTTTGTGCTGCTGATGTTCAACACCTGGTACACAAACCATCGGAGAGA 1427
Qy 749 GGGTCAGGAGATCAATTTGATGCTGTCAAAATTTGAGCAGGAGTTTAAACAGAACCT 808
Db 1426 GAGTAAGAGAAATAATTTATCAATGCTGTTCCGATAGAACAGGAGTTCTCTCAGGSCCT 1367
Qy 809 TGCAGTTGGCTCAATTTGGAATGAAATGCAATTTGATGAAACAGATACATGAGTTTGTAG 868
Db 1366 TGCTGTGAAGCTCAATTTGGGATGAATTTGCACTCTAATGAAACATACATGAGTTTGTGG 1307
Qy 869 CTGACAGATTAATTTGTGGAATCTGATTTCTCAAGGTTTTCAGGCAAGAAATCCTTTTG 928
Db 1306 CAGACAGATTAATTTGCTGGAATCTGATTTTGGATGAAACAGATACATGAGTTTGTAG 1247
Qy 929 ATTTTATGAAAACATTTCTTTAGAGGAAAACAAATTTCTTTGAGAAAACGAGTTTTCAG 988
Db 1246 ACTTATGAGAAATATTTCTCTGGAAGAAAGACTAATCTTTTGAAGAGAGATAGGCG 1187
Qy 989 AGTATCAGCGTTTTCAGTTATGCGAGAAACAGATACAGTCTTCCACCTTGGATGAG 1048
Db 1186 AGTATCAGAGATGGGAGTGTATGTCAGATCCACAGAGAAATCTTTTACCTTGGATGCTG 1127
Qy 1049 ATTT 1052
Db 1126 ACTT 1123

RESULT 13

ABL65414
ID ABL65414 standard; DNA; 2500 BP.

XX ABL65414;

XX AC

XX 15-MAY-2002 (first entry)

XX Lung cancer related gene sequence SEQ ID NO:3751.

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.

XX Homo sapiens.
OS WO200194629-A2.
PN 13-DEC-2001.
PD 30-MAY-2001; 2001WO-US010838.
XX 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
XX 18-SEP-2000; 2000US-0233133P.
PR 18-SEP-2000; 2000US-0233617P.
XX 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
XX 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
XX 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
XX 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
XX 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
XX 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
XX 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
XX 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
XX 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
XX 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
XX 28-SEP-2000; 2000US-0236109P.
PR 29-SEP-2000; 2000US-0236111P.
XX 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
XX 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
XX 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
XX 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
XX 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
XX 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
XX 03-OCT-2000; 2000US-0237606P.
PR 01-NOV-2000; 2000US-0244867P.
XX 01-NOV-2000; 2000US-0245084P.
PA (AVAL-) AVALON PHARM.
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX WPI; 2002-188264/24.

Screening for anti-neoplastic agent involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, and determining a change in expression of a gene of a signature gene set.

Claim 1; SEQ ID NO 3751; 44pp; English.

The present invention describes a method (M1) for screening for an anti-neoplastic agent. The method involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, determining a change in expression of at least one gene (I) of a signature gene set, where (I) comprises a sequence (S) selected from 8447 sequences (given in ABL61664 to ABL70110), or is at least 95% identical to (S), where a change in expression is indicative of anti-neoplastic activity. (I) has cytostatic activity and can be used in gene therapy. M1 can be used for screening an

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2005, 22:53:45 ; Search time 2718.5 Seconds
(without alignments)
14744.041 Million cell updates/sec

Title: US-10-698-228-2

Perfect score: 1053

Sequence: 1 atggcgaccggaaagcc.....tcacctggatcagatcttt 1053

Scoring table: IDENTITY_NUC

Gapop 10_0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gss1:*

9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1053	100.0	3292	CR617553	full-leng
2	1053	100.0	4748	AL137348	Homo sapi
3	1041	98.9	4650	BC042948	Homo sapi
4	1010.6	96.0	3501	CR627376	Homo sapi
5	945	89.7	1056	AY398973	Homo sapi
6	929.6	88.3	1083	AL547501	AL547501
7	829.8	78.8	1008	AY398974	Pan. trogl
8	793.2	75.3	879	BQ441857	AGENCOURT
9	777	73.8	1043	AY398975	Mus muscu
10	731.2	69.4	892	CNI63214	952695 MA
11	728.2	69.2	914	CF995079	AGENCOURT
12	710.6	67.5	870	CO648380	ILLUMIGEN
13	682.8	64.8	804	CD566661	AGENCOURT
14	655.4	62.2	1061	BM801298	AGENCOURT
15	654.8	62.2	896	BUI70979	AGENCOURT
16	616.4	58.5	685	BX951374	DKFZp781E
17	578.4	54.9	1571	CR603461	full-leng
18	578.4	54.9	1573	CR625489	full-leng
19	578.4	54.9	1582	CR608076	full-leng
20	578.4	54.9	1588	CR602054	full-leng
21	578.4	54.9	1592	CR621427	full-leng
22	578.4	54.9	1600	CR604378	full-leng
23	578.4	54.9	1605	CR590959	full-leng
24	578.4	54.9	1612	CR614990	full-leng

25	578.4	54.9	1613	3	CR609838	full-leng
26	578.4	54.9	1623	3	CR618451	full-leng
27	578.4	54.9	1630	3	CR602150	full-leng
28	578.4	54.9	1796	3	CR596700	full-leng
29	570.4	54.2	2088	3	AK088907	Mus muscu
30	562	53.4	638	7	CN358022	170005831
31	547	51.9	547	7	CN358025	170005326
32	530.8	50.4	836	7	CR416843	CR416843
33	528.8	50.2	608	1	AU137050	AU137050
34	527.2	50.1	540	5	BUI29508	UI-HF-BNO
35	525	49.9	997	4	BM468712	AGENCOURT
36	518.8	49.3	522	5	BQ694828	1000971 H
37	513.6	48.8	793	7	CK778582	965745 MA
38	503.8	47.8	923	7	CR580780	CR580780
39	503	47.8	792	5	BUI28458	603494277
40	502	47.7	510	7	CR540284	DKFZp459N
41	501.4	47.6	697	6	CA328968	UI-M-FY0-
42	493.8	46.9	510	5	BX955728	DKFZp781G
43	493.2	46.8	946	7	CN024619	AGENCOURT
44	491.8	46.7	932	5	BUI96941	AGENCOURT
45	489.4	46.5	909	6	CA981614	AGENCOURT

ALIGNMENTS

RESULT 1

CR617553

LOCUS

DEFINITION

of Homo sapiens (human).

ACCESSION

CR617553

VERSION

CR617553.1

KEYWORDS

HTC; CNSLT_CDNA.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 3292)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished

JOURNAL

REMARK

Contact : Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Faraday Avenue

Genoscope.

REFERENCE

2 (bases 1 to 3292)

Direct Submission

Submitted (20-JUL-2004)

Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library

was normalized. Library was constructed by Life Technologies, a

division of Invitrogen.

Location/Qualifiers

1. 3292

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0D1011YF14"

/tissue_type="Placenta Cot 25-normalized"

/plasmid="pCMVSPORT_6"

ORIGIN

Query Match 100.0%; Score 1053; DB 3; Length 3292;

Best Local Similarity 100.0%; Pred. No. 3.3e-269;

Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGGCGACCCGGAAGCGCGGCGGCTGGATCAGGATCAGATCATCTTCA 60

|||||

27 ATGGGCGACCCGGAAGCGCGGCGGCTGGATCAGGATCAGATCATCTTCA 86


```
QY 181 GCTTCCTCTCGACGACGAGAGGTCGACTTATCAAGGATCTCCCTCACTCGGAACAAG 240
Db |||
QY 227 GCTTCCTCTCGACGACGAGAGGTCGACTTATCAAGGATCTCCCTCACTCGGAACAAG 286
Db |||
QY 241 CTTAAAGCAGATGAGAGTACTTCATCTCTCACTCTTACCTTTTTCAGCCAGTGTAT 300
Db |||
QY 287 CTTAAAGCAGATGAGAGTACTTCATCTCTCACTCTTACCTTTTTCAGCCAGTGTAT 346
QY 301 GGAATGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
Db |||
QY 347 GGAATGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 406
QY 361 CCCTGTTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGACAGTTTG 420
Db |||
QY 407 CCCTGTTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGACAGTTTG 466
QY 421 CTGATAGACACTTATCATCAGAGATCCCAAGAAAAGGAAATTTTATTTAATGCAATTGAA 480
Db |||
QY 467 CTGATAGACACTTATCATCAGAGATCCCAAGAAAAGGAAATTTTATTTAATGCAATTGAA 526
QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATGGGCTTCGGATGGATAGCAGATAGAAA 540
Db |||
QY 527 ACCATGCCCTATGTTAAGAAAAGCAGATGGGCTTCGGATGGATAGCAGATAGAAA 586
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGTGTAGAGAGGATTTTCTTCTCAGGA 600
Db |||
QY 587 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGTGTAGAGAGGATTTTCTTCTCAGGA 646
QY 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCGAGACTCACTTTTCC 660
Db |||
QY 647 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCGAGACTCACTTTTCC 706
QY 661 AATGAACATCATCAGCAGAGATGAAGACTTCTAGTACTTTGCTTGGCTGATGTTCCAA 720
Db |||
QY 707 AATGAACATCATCAGCAGAGATGAAGACTTCTAGTACTTTGCTTGGCTGATGTTCCAA 766
QY 721 TACTTAGTAAATAAGCCTTCAGAGAAAGGTCAGGAGATCATTTGTCATCTGTCAAA 780
Db |||
QY 767 TACTTAGTAAATAAGCCTTCAGAGAAAGGTCAGGAGATCATTTGTCATCTGTCAAA 826
QY 781 ATTGAGCAGGAGTGTAAACAGAGCCTTGCAGTTGGCTCATTTGGAATGAATTCATT 840
Db |||
QY 827 ATTGAGCAGGAGTGTAAACAGAGCCTTGCAGTTGGCTCATTTGGAATGAATTCATT 886
QY 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
Db |||
QY 887 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 946
QY 901 AAGTTTTTTCAGGCGAGAAAATCCTTTGATTTTATGGAAGAAACATTTCTTTAGAGGAAA 960
Db |||
QY 947 AAGTTTTTTCAGGCGAGAAAATCCTTTGATTTTATGGAAGAAACATTTCTTTAGAGGAAA 1006
QY 961 ACAAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTATGSCAGAAACC 1020
Db |||
QY 1007 ACAAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTATGSCAGAAACC 1066
QY 1021 ACAGATAACGTTTTCACCTTTGGATGACAGATTTT 1053
Db |||
QY 1067 ACAGATAACGTTTTCACCTTTGGATGACAGATTTT 1099
Db |||

RESULT 3
BC042948
LOCUS Homo sapiens cdna clone IMAGE:4798175, containing frame-shift errors.
ACCESSION BC042948
VERSION BC042948.2 GI:34194000
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 4650)
Strausberg R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Udwin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Maman, A., Rodriguez, S.,
Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 4650)
Director MGC Project.
Direct Submission
Submitted (09-JAN-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgi.nci.nih.gov
On Aug 25, 2003 this sequence version replaced gi:27695575. ---
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth
Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prabh,
Parvaneh Saeedi, JR Santos, Angelique Schnerch, Ursula Skalska,
Duane Smal, Jeff Stott, Miranda Tsai, George Yang, Jacque
Schein, Asim Siddiqui, Rob Holt, Marco Marra.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 70 Row: h Column: 19
This clone has the following problem: frame shifted.
FEATURES
source
1. 4650
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4798175"
/tissue_type="Brain, hypothalamus"
/clone_lib="NIH MGC_96"
/lab_host="DH10E"
/notes="Vector: pBluescript"
```

ORIGIN

Query Match 98.9%; Score 1041; DB 3; Length 4650;
Best Local Similarity 99.9%; Pred. No. 5.8e-266;
Matches 1052; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY

1 ATGGGCGACCCGGAAGCGCGGAGCGCGCGGCTGGATCAGATGAGATCATCTTCA 60
|||||

Db	72	ATGGGCGACCCGGAAAGCGCGCGGCTGGATCAGAGTGAAGATCATCTTCA	131
QY	61	GACACAAACGAAAGTGAATAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTCTCTCG	120
Db	132	GACACCAACGAAAGTGAATAAGTCAAAATGAAGAG-CACTCTAAGAAAGAGTCTCTCG	190
QY	121	CGTTTGTGCATCTTTCCAAATCCAGTACCCTGTATTTTGGAAAATGATPATAACACGAGCAG	180
Db	191	CGTTTGTGCATCTTTCCAAATCCAGTACCCTGTATTTTGGAAAATGATPATAACACGAGCAG	250
QY	181	GCTTCTCTTGACAGCAGACGAGGTGCGATTCATCAAGGATCTCCCTCACTGGAAACAG	240
Db	251	GCTTCTCTTGACAGCAGACGAGGTGCGATTCATCAAGGATCTCCCTCACTGGAAACAG	310
QY	241	CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCAATCTTAGCCCTTTTTCGAGCCAGTGAT	300
Db	311	CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCAATCTTAGCCCTTTTTCGAGCCAGTGAT	370
QY	301	GGAAATGTAAATGAAATTTGGTGGAGCGCTTTTAGTCAGAGGTGCAGGTTCAGAGGCT	360
Db	371	GGAAATGTAAATGAAATTTGGTGGAGCGCTTTTAGTCAGAGGTGCAGGTTCAGAGGCT	430
QY	361	CGCTGTTCTATGGCTTTCAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG	420
Db	431	CGCTGTTCTATGGCTTTCAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG	490
QY	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA	480
Db	491	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA	550
QY	481	ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGATGGATAGCAGATAGAAA	540
Db	551	ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGATGGATAGCAGATAGAAA	610
QY	541	TCATCTTTTGGGAAAGAGTGGTGGCTTTCGCTGTGAAGAGAGTTCCTCTCAGGA	600
Db	611	TCATCTTTTGGGAAAGAGTGGTGGCTTTCGCTGTGAAGAGAGTTCCTCTCAGGA	670
QY	601	TCCTTTTGTGCTATATTCCTGGCTTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTCC	660
Db	671	TCCTTTTGTGCTATATTCCTGGCTTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTCC	730
QY	661	AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTTGCCTTGCCTGATGTTCCAA	720
Db	731	AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTTGCCTTGCCTGATGTTCCAA	790
QY	721	TACTTAGTAAATAGCCTTCAGAAAGAGGTGAGGAGATCATTTGTTGATGCTGTCAA	780
Db	791	TACTTAGTAAATAGCCTTCAGAAAGAGGTGAGGAGATCATTTGTTGATGCTGTCAA	850
QY	781	ATTGAGCAGGAGTCTTTTACAGAGGCTTGGCAGTTGGCTCATTTGGAATGAATTCATTT	840
Db	851	ATTGAGCAGGAGTCTTTTACAGAGGCTTGGCAGTTGGCTCATTTGGAATGAATTCATTT	910
QY	841	TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTTGTGGAACTTGGATTTCA	900
Db	911	TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTTGTGGAACTTGGATTTCA	970
QY	901	AAGGTTTTTCAGGAGAAAAATCCTTTTGAATTTTATGGAATAAATTTCTTTTAGAAGAAAA	960
Db	971	AAGGTTTTTCAGGAGAAAAATCCTTTTGAATTTTATGGAATAAATTTCTTTTAGAAGAAAA	1030
QY	961	ACAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTATTGCGAGAAACC	1020
Db	1031	ACAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTATTGCGAGAAACC	1090
QY	1021	ACAGATAACGCTTCTACCTTTGGATGCAGATTTT	1053
Db	1091	ACAGATAACGCTTCTACCTTTGGATGCAGATTTT	1123

LOCUS	CR627376	3501 bp	mRNA	linear	HTC 22-SEP-2004
DEFINITION	Homo sapiens mRNA; cDNA DKFZp686M05248 (from clone DKFZp686M05248).				
ACCESSION	CR627376				
VERSION	CR627376.1	GI:50949847			
KEYWORDS	HTC.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 3501)				
AUTHORS	Ottenwaelder,B., Obermaier,B., Deutschenbaur,S., Schaipp,A., Mewes,H.W., Weill,B., Amdt,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.				
CONSRTM	The German cDNA Consortium				
TITLE	Direct Submission				
JOURNAL	Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764				
COMMENT	Neuherberg, GERMANY Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by Medigenomix (Martinsried/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp686M05248) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering: http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp686M05248 Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.				
FEATURES	Location/Qualifiers				
source	1..3501				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/clone="DKFZp686M05248"				
	/tissue_type="salivary gland"				
	/clone_lib="686 (synonym: hlccc3). Vector pSport1_Sfi; host DH10B; sites SfiIA + SfiIB"				
	/dev_stage="adult"				
	/note="ribonucleotide reductase M2 B (TP53 inducible)"				
gene	1..3501				
	/gene="DKFZp686M05248"				
CDS	332..1225				
	/gene="DKFZp686M05248"				
	/codon_start=1				
	/product="hypothetical protein"				
	/protein_id="CAH10473.1"				
	/db_xref="GI:50949848"				
	/translation="MYKQAQASFWTAEVDLSKDLPHWNKKADEKFIYSHILARPPAASDGIYVNNLVERFSQVQVPEARCFYGFQILINHVSEMYSLIDITYIRDPKRRBFLPNATETPPYKKKADWALRWIADRKSTFGERVAFVAAVEGVFGSFAAIFWLKKRGLMPLGTSNELISRDEGLHCDPACIMFQYLVNKPPEERVREIIVDAVKLEQFLETLALPVGLIGMNCILMKQYIEFVADRLLVELGFSKVFOAENPFDPMENISLEKTFNPFEXRVS YQRFVAVMAETTDNVFTLDADF"				
ORIGIN					
Query Match	96.0%; Score 1010.6; DB 3; Length 3501;				
Best Local Similarity	98.6%; Pred. No. 6.8e-258;				
Matches 1019; Conservative	0; Mismatches 14; Indels 0; Gaps 0;				
QY	21	GGAAGCGCGCGGCTGGATCAGGATGAGAGATCATCTTCAGACACCAACGAAAGTGAAT	80		
Db	190	GGACGACGCTTGGAGGCTTCAGCGCGGAGATCATCTTCAGACACCAACGAAAGTGAAT	249		
QY	81	AAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGCGGGTTGTCTATCTTTCCAAAT	140		
Db	250	AAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGCGGGTTGTCTATCTTTCCAAAT	309		
QY	141	CCAGTACCTCGATATTTGGAAAAATGTATAAACAGGCACAGGCTTCCTCTTCGACAGCAGA	200		
Db	310	CCAGTACCTCGATATTTGGAAAAATGTATAAACAGGCACAGGCTTCCTCTTCGACAGCAGA	369		
QY	201	AGAGGTGCGACTTATCAAGGATCTCCCTCACTGGAAACAGCTTTAAAGCAGATGAGAAGTA	260		
Db	370	AGAGGTGCGACTTATCAAGGATCTCCCTCACTGGAAACAGCTTTAAAGCAGATGAGAAGTA	429		

QY 261 CTTCACTCTCACATCTTAGCCCTTTTTCAGCCAGTGATGGAATTTGTAATGAAAATTT 320
Db |||||
QY 430 CTTCACTCTCACATCTTAGCCCTTTTTCAGCCAGTGATGGAATTTGTAATGAAAATTT 489
Db |||||
QY 321 GGTGGAGCCCTTTAGTCAGAGGTGAGGTTCCAGAGGTCGCTGTTTCTATGCGCTTCA 380
Db |||||
QY 490 GGTGGAAACCTTTAGTCAGAGGTGAGGTTCCAGAGGTCGCTGTTTCTATGCGCTTCA 549
Db |||||
QY 381 AATTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTGCTGATAGACACTTACATCAG 440
Db |||||
QY 550 AATTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTGCTGATAGACACTTACATCAG 609
Db |||||
QY 441 AGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTTGAAACCATGCCCTATGTTAAGAA 500
Db |||||
QY 610 AGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTTGAAACCATGCCCTATGTTAAGAA 569
Db |||||
QY 501 AAAAGCAGATTTGGCCCTTCGATGGATAGCAGATAGAAAATCTACTTTTGGGGAAGAGT 560
Db |||||
QY 670 AAAAGCAGATTTGGCCCTTCGATGGATAGCAGATAGAAAATCTACTTTTGGGGAAGAGT 729
Db |||||
QY 561 GGTGGCCTTTGCTGCTAGAGAGGTTTCTTCTCAGGATCTTTTGTGCTATATTCG 620
Db |||||
QY 730 GGTGGCCTTTGCTGCTAGAGAGGTTTCTTCTCAGGATCTTTTGTGCTATATTCG 789
Db |||||
QY 621 GCTAAAGAAAGAGAGGTCTTATGCCAGGACTCACTTTTCCAAATGAATCATCAGCAGAGA 680
Db |||||
QY 790 GCTAAGAAAGAGAGGTCTTATGCCAGGACTCACTTTTCCAAATGAATCATCAGCAGAGA 849
Db |||||
QY 681 TGAAGCACTTCACTGTGACTTTTCTTGTGCTGATGTTTCCAAATGAATTAAGCCTTC 740
Db |||||
QY 850 TGAAGCACTTCACTGTGACTTTTCTTGTGCTGATGTTTCCAAATGAATTAAGCCTTC 909
Db |||||
QY 741 AGAAGAAAGGTCAGGAGATCATGTTGATGCTGTCAAAAATTGAGCAGAGTTTTTTAAC 800
Db |||||
QY 910 AGAAGAAAGGTCAGGAGATCATGTTGATGCTGTCAAAAATTGAGCAGAGTTTTTTAAC 969
Db |||||
QY 801 AGAAGCCTTGCAGTTGGCTCATTTGGAATGAATTCATTTTGGTGAACACAGTACATCA 860
Db |||||
QY 970 AGAAGCCTTGCAGTTGGCTCATTTGGAATGAATTCATTTTGGTGAACACAGTACATCA 1029
Db |||||
QY 861 GTTTGTAGCTGACAGATTAATTTGTGAACTTGGATTTCAAAAGTTTTTTCAGCAGAAAA 920
Db |||||
QY 1030 GTTTGTAGCTGACAGATTAATTTGTGAACTTGGATTTCAAAAGTTTTTTCAGCAGAAAA 1089
Db |||||
QY 921 TCTTTTGAATTTAGGAAAAATTTCTTTTGAAGAAAAAACAATTTCTTTGAGAAACG 980
Db |||||
QY 1090 TCTTTTGAATTTAGGAAAAATTTCTTTTGAAGAAAAAACAATTTCTTTGAGAAACG 1149
Db |||||
QY 981 AGTTTCAGATATCAGCGTTTTTGCAGTTATGCGAAGAACCCACAGATAAGCTTTCACCTT 1040
Db |||||
QY 1150 AGTTTCAGATATCAGCGTTTTTGCAGTTATGCGAAGAACCCACAGATAAGCTTTCACCTT 1209
Db |||||
QY 1041 GGATGCAGATTTT 1053
Db |||||
QY 1210 GGATGCAGATTTT 1222
Db |||||

RESULT 5
AY398973
LOCUS
DEFINITION Homo sapiens HCM0069 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION
VERSION AY398973.1 GI:39754962
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1056)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,

Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios
 Science 302 (5652), 1960-1963 (2003)
 JOURNAL 14671302
 PUBMED 2 (bases 1 to 1056)
 REFERENCE
 AUTHORS Clark,A.G., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 Direct Submission
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES
 Location/Qualifiers
 source 1..1056
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 <1..>1056
 /locus_tag="HCM0069"

Query Match 89.7%; Score 945; DB 9; Length 1056;
Best Local Similarity 89.7%; Pred. No. 1.5e-240;
Matches 945; Conservative 0; Mismatches 108; Indels 0; Gaps 0;
QY 1 ATGGGGCAGCCCGGAAAGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db |||||
QY 61 GACACAAAGAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTGC 120
Db |||||
QY 121 CGGTTTGTCTATCTTTCCAAATCCAGTACCCTGATATTTGGAAATGTATAACAGGCACAG 180
Db |||||
QY 181 GCTTCCTCTTGACAGCAGAGAGGTGCGACTTATCAAGAGATCTCCCTCACTGGAACAAG 240
Db |||||
QY 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCACATCTTAGCCCTTTTTCAGCCAGTGAT 300
Db |||||
QY 301 GGAAATGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGCT 360
Db |||||
QY 361 CGCTGTTTCTATGGCTTCAAAATTTCTCATCGAGAATGTTTCACTCAGAGATGTACAGTTTG 420
Db |||||
QY 421 CTGATAGACACTTATACATCAGAGATCCCAAGAAAGGGAATTTTTTATTAATGCAATTGAA 480
Db |||||
QY 481 ACCATGCCCTATGTAAAGAAAAAGCAGATTTGGCGCTTGGCATGGATAGCAGATAGAAAA 540
Db |||||
QY 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTAGAGAGAGTTTCTTCTCAGGA 600
Db |||||
QY 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAAAGAGAGGTCTTATGCCAGGACTCCTTTTTC 660
Db |||||
QY 660 TCTTTTGTGCTATATTTCTGGCTTAAAGAAAGAGAGGTCTTATGCCAGGACTCCTTTTTC 660
Db |||||


```
QY 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATGTGGCCCTTGGATGGATGAGATAGAGAA 540
Db 481 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 540
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTGTGCTGTAGAGAGCTTTCTCTCAGGA 600
Db 541 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 600
QY 601 TCTTTTGTCTATATCTTGGCTTAAAGAAAGAGAGGTCTTATGCCAGGACTCACTTTTCC 660
Db 601 TCAATTGTGCAATATCTTGGCTTAAAGAAAGAGAGGCTCATGCTGGAGTCACTTTTCA 660
QY 661 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA 720
Db 661 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA 720
QY 721 TACTTGTAGTAATAGCTTTCAGAAAGAGGTTCAGGAGATCATTTGTGATGCTGTCAA 780
Db 721 TACTTGTAGTAATAGCTTTCAGAAAGAGGTTCAGGAGATCATTTGTGATGCTGTCAA 780
QY 781 ATTGACGAGGAGTTTAAACAGAGCTTGCAGTTGGCTCATTTGGAATGAATTCATT 840
Db 781 ATCGAG-----CAGAGCTTGGCTTGGCTCATTTGGAATGAATTCATT 840
QY 841 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTAATTTGTGGAATTTGGAATTCAT 900
Db 841 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTAATTTGTGGAATTTGGAATTCAT 900
QY 901 AAGGTTTTTTCAGCAGAGAAATCCTTTTGAATTTATGGAAGAAATTTCTTTTGAAGAA 960
Db 901 AAGGTTTTTTCAGCAGAGAAATCCTTTTGAATTTATGGAAGAAATTTCTTTTGAAGAA 960
QY 961 ACATATTTCTTTCAGAAAGAGTTTCAGAGATCATCAGCTTTTGCAGTTATGCGAGAAC 1020
Db 961 ACATATTTCTTTCAGAAAGAGTTTCAGAGATCATCAGCTTTTGCAGTTATGCGAGAAC 1020
QY 1021 ACAGATAACGCTTTCACCTTTGGATGAGATTT 1052
Db 1008 ACAGATAACGCTTTCACCTTTGGATGAGATTT 1039
```

```
RESULT 10
LOCUS CN163214/c 892 bp mRNA linear EST 02-APR-2004
DEFINITION 952895 MARC 4P1G Sus scrofa cDNA 3', mRNA sequence.
ACCESSION CN163214
VERSION CN163214.1 GI:46177644
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 892)
Smith, T.P.L., Freking, B.A., Ford, J.J., Vallet, J.L., Wise, T.A.,
Noneman, D.J., Wray, J.E. and Keele, J.W.
Porcine EST collection using a normalized library constructed from
embryos representing early developmental stages
Unpublished (2003)
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross_match v0.990329.
Place: TMR8060 row: G column: 4
Seq primer: TAGAAGGCACAGTCGAGG.
1. .892
Location/Qualifiers
/organism="Sus scrofa"
/mol_type="mRNA"
```

FEATURES

source

RESULT 11
CF995079
LOCUS

CF995079 914 bp mRNA linear EST 25-NOV-2003

```
/db_xref="taxon:9823"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 4P1G"
/notes="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI;
Library made with combined RNA from day-10, day-13,
day-15, day-25, and day-30 whole embryos."
```

ORIGIN

```
Query Match 69.4%; Score 731.2; DB 7; Length 892;
Best Local Similarity 94.1%; Pred. No. 1.4e-183;
Matches 760; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 246 AGCAGATGAGAAGTACTTCTCATCTCTACATCTTAGCCCTTTTTCAGCCAGTGTGAAT 305
Db 892 ATCAGATGAGAGTATTTATCTCTCATCTTAGCCCTTTTTCAGCCAGTGTGAAT 833
QY 306 TGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGTTCCAGAGGCTCGCTG 365
Db 832 TGTGAATGAAACCTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGTTCCAGAGGCGCTG 773
QY 366 TTTCTATGGCTTTCAAAATTTCTCATCGAATGTTCACTCAGAGATGTACAGTTTGTGAT 425
Db 772 TTTCTACGGCTTTCAAAATTTCTCATCGAATGTTCACTCAGAGATGTACAGTTTGTGAT 713
QY 426 AGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATCAATTTGAAACCAT 485
Db 712 AGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATCAATTTGAGACAT 653
QY 486 GCCTATGTTTAAAGAAAAAGCAGATGGGCTTTCGATGATAGCAGATGAAATCTTAC 545
Db 652 GCCATATGTTAAGAAAAAGCAGATGGGCTTTCGATGATAGCAGATGAAAGCTTAC 593
QY 546 TTTTGGGAAAAGTGGTGGCTTTTGTCTGTGTAGAAAGAGTTTCTTCTCAGAGATCTTT 605
Db 592 TTTTGGGAAAAGTGGTGGCTTTTGTCTGTGTAGAAAGAGTTTCTTCTCAGAGATCTTT 533
QY 606 TGCTGCTATATCTGGCTTAAAGAAAGAGAGCTTATGCCAGGACTCACTTTTCCAAATGA 665
Db 532 TGCTGCTATATCTGGCTTAAAGAAAGAGAGCTTATGCCAGGACTCACTTTTCCAAATGA 473
QY 666 ACTCATCAGCAGAGATGAAAGGACTTCACTGTGACTTTTGTCTGTGATGTTTCCAAATCTT 725
Db 472 ACTCATCAGCAGAGATGAAAGGACTTCACTGTGACTTTTGTCTGTGATGTTTCCAAATCTT 413
QY 726 AGTAAATAAGCTTTCAGAAAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAAAATGA 785
Db 412 GGTAAATAAGCTTTCAGAAAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAAAATGA 353
QY 786 GCAGGAGTTTAAACAGAGACCTTGCAGTTGGCTCATTTGGAATGAATTTGCATTTTGTAT 845
Db 352 GCAGGAGTTTAAACAGAGACCTTGCAGTTGGCTCATTTGGAATGAATTTGTTGTTTGTAT 293
QY 846 GAAACAGTACATTGAGTTTGTAGCTGACAGATTAATTTGTGAACTTCGATTTCTCAAGGT 905
Db 292 GAAACAGTATATTGAGTTTGTAGCTGACAGATTAATTTGTGAACTTCGATTTCTCAAGGT 233
QY 906 TTTTTCAGGAGAAATTCCTTTTGAATTTTATGGAAGAAATTTCTTTTGAAGGAAACAAA 965
Db 232 TTTTTCAGGAGAAATTCCTTTTGAATTTTATGGAAGAAATTTCTTTTGAAGGAAACAAA 173
QY 966 TTTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTCAGGTTTATGCGAGAAACACAGA 1025
Db 172 TTTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTCAGGTTTATGCGAGAAACACAGA 113
QY 1026 TAACGCTTTCACCTTCGATGAGATTTT 1053
Db 112 TAATGCTTTCACCTTCGATGAGATTTT 85
```



```
/clone="IBIUW:25149"
/sex="male"
/tissue_type="blood"
/cell_type="PBMC"
/dev_stage="adult"
/lab_host="Electronmax DH10B"
/clone_lib="Katz MPB2"
/notes="Vector: pDONR 222; Site 1: BarG I; Site 2: BarG I;
Created from CloneMiner cDNA Library Construction kit
(catalog #18249-029)"

ORIGIN
Query Match 67.5%; Score 710.6; DB 7; Length 870;
Best Local Similarity 98.8%; Pred. No. 4.4e-178;
Matches 716; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 329 GCTTTAGTCAGAGGTCAGAGGTCAGAGGTCGCTGTTCTATGCGCTTCAAAATCTCA 388
Db |||
QY 389 TCGAGAAATCTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAGATCCCA 448
Db |||
QY 63 TCGAGAAATCTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAGATCCCA 122
Db |||
QY 449 AGAAAGGGAATTTTATTAATGCAATTCAAACCATGCCCTATGTTAAGAAAGGAGCAG 508
Db |||
QY 123 AGAAAGGGAATTTTATTAATGCAATTCAAACCATGCCCTATGTTAAGAAAGGAGCAG 182
Db |||
QY 509 ATTGGGCTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 568
Db |||
QY 183 ATTGGGCTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 242
Db |||
QY 569 TTGCTGCTGTAAGAGGATTTTCTCTCAGGATCTTTCTGCTGCTGCTGCTGCTGCTGCTGCT 628
Db |||
QY 243 TTGCTGCTGTAAGAGGATTTTCTCTCAGGATCTTTCTGCTGCTGCTGCTGCTGCTGCTGCT 302
Db |||
QY 629 AGAGAGGCTTATGCCAGGACTCACATTTTCCAAATGCAATTCATCAGCAGATGAAGGAC 688
Db |||
QY 303 AGAGAGGCTTATGCCAGGACTCACATTTTCCAAATGCAATTCATCAGCAGATGAAGGAC 362
Db |||
QY 689 TTCACGTGACTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 748
Db |||
QY 363 TTCACGTGACTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 422
Db |||
QY 749 GGGTCAGGAGATCATTTGATGCTGCTCAAAATTCAGCAGAGATTTTAAACAGAGCCT 808
Db |||
QY 423 GGGTCAGGAGATCATTTGATGCTGCTCAAAATTCAGCAGAGATTTTAAACAGAGCCT 482
Db |||
QY 809 TGCCAGTTGGCCTCATTTGGAATGAATTTGATTTGATGAACAGTACATTTGAGTTGTAG 868
Db |||
QY 483 TGCCAGTTGGCCTCATTTGGAATGAATTTGATTTGATGAACAGTACATTTGAGTTGTAG 542
Db |||
QY 869 CTGACAGATTAATCTTGGAATCTGGAATCTCAAGGTTTTCAGGAGAGAAATCCCTTTTG 928
Db |||
QY 543 CTGACAGATTAATCTTGGAATCTGGAATCTCAAGGTTTTCAGGAGAGAAATCCCTTTTG 602
Db |||
QY 929 ATTTATGAAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACAGATTTTCAG 988
Db |||
QY 603 ATTTATGAAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACAGATTTTCAG 662
Db |||
QY 989 AGTATCAGGTTTTCAGTATGAGGAGAAACACAGATACCTCTTCACTTTGGATGAG 1048
Db |||
QY 663 AGTATCAGGTTTTCAGTATGAGGAGAAACACAGATACCTCTTCACTTTGGATGAG 722
Db |||
QY 1049 ATTTT 1053
Db |||
QY 723 ATTTT 727
Db |||
```

RESULT 13

CD656661 804 bp mRNA linear EST 18-JUN-2003
LOCUS AGENCOURT_14555875 NIA Human H1 Embryonic Stem Cell cDNA Library

ORIGIN

Query Match 64.8%; Score 682.8; DB 6; Length 804;
Best Local Similarity 97.3%; Pred. No. 1.1e-170;
Matches 716; Conservative 0; Mismatches 17; Indels 3; Gaps 2;

(long) Homo sapiens cDNA clone IMAGE:30427288 5', mRNA sequence.
CD656661 GI:31897180
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 804)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabs-remail.nih.gov
Tissue Procurement: Irene Ginis and Mahendra Rao, NIA
cDNA Library Preparation: Yulan Piao and Minoru Ko
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC c lone distribution information
can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: NDAM514 row: h column: 17
High quality sequence stop: 680.
Location/Qualifiers
1. 804
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30427288"
/tissue_type="Embryonic Stem cells"
/cell_line="WA01"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIA Human H1 Embryonic Stem Cell cDNA Library (Long)"
/note="Vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI;
This is a long-transcript enriched cDNA library (Genome
Res. 11: 1553-1558 (2001). [PMID: 11544199]) from WA01
cell line. Undifferentiated human ES cell line WA01/H1
was obtained from WiCell Research Institute, Inc.,
Madison, WI, cultured according to their instructions, on
MEF feeders. They formed round colonies with defined edges
and were positive for alkaline phosphatase, SSEA-4, OCT3,
OCT4, REX1, UTP, TERT, SOX2, CX43 and CX45. They are
negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLT3, SSEA-1,
TUBB3, NES, GFAP, and BOMES. When confluent (18-10 days
after plating), the ES cells from 4 x 6cm dishes were
treated with 1 mg/ml collagenase, type IV
(Invitrogen/GIBCO) for 5-10 min and gently scraped off
with 5 ml pipette. RNA was purified with Trizol Reagent
from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558
(2001). [PMID:11544199]) Double-stranded cDNAs were
synthesized with an Oligo(dT) primer [Invitrogen:
5'-pGACTAGTCTAGTCGAGCGCCCTTTTCTTTTCTTTT-3'] from
3.4g of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to Loner-linker IL-Sal4, purified by phenol/chloroform
extraction, and separated from free linkers by
Centricon-100 column. Then, the cDNAs were amplified by
long-range high fidelity PCR using Ex Taq polymerase
(Takara) with a primer Sal4-S for 25 cycles. The products
were purified by phenol/chloroform extraction and
Centricon-100 column. The cDNAs were digested with SalI
and NotI enzymes and cloned into SalI/NotI site of
pCMV-SPORT6 plasmid vector. The average insert size is
about 3.6kb."

```
QY 212 TATCAAGGATCTCCCTCACTGGAAACAAGCTTTAAAGCAGATGAGAAGTACTTTCATCTCTC 271
Db 1 TATCAAGGATCTCCCTCACTGGAAACAAGCTTTAAAGCAGATGAGAAGTACTTTCATCTCTC 60
QY 272 ACATCTTAGC-CTTTTTTTCAGCCAGTGTAAATGTAATGAAATTTGGTGGAGCGC 330
Db 61 ACATCTTAGCNCCTTTTTTTCAGCCAGTGTAAATGTAATGAAATTTGGTGGAGCGC 120
QY 331 TTTAGTCAGGAGTGCAGGTTCCAGAGGCTCCCTGTTTCTATGGCTTTCAAAATTTCTCATC 390
Db 121 TTTAGTCAGGAGTGCAGGTTCCAGAGGCTCCCTGTTTCTATGGCTTTCAAAATTTCTCATC 180
QY 391 GAGATGTTCACTCAGAGATGTACAGTTGCTGTAGACACTTACATCAGAGATCCCAAG 450
Db 181 GAGATGTTCACTCAGAGATGTACAGTTGCTGTAGACACTTACATCAGAGATCCCAAG 240
QY 451 AAAAGGGAATTTTTTAAATGCAATGTAAGCAATGCTCCCTATGTTAAAGAAAAAGCAGAT 510
Db 241 AAAAGGGAATTTTTTAAATGCAATGTAAGCAATGCTCCCTATGTTAAAGAAAAAGCAGAT 300
QY 511 TGGGCTTTCGATGGATAGACAGATAAATCTACTTTTGGGGAAGAGTGGTGGCTTT 570
Db 301 TGGGCTTTCGATGGATAGACAGATAAATCTACTTTTGGGGAAGAGTGGTGGCTTT 360
QY 571 GCTGCTGTAGAAAGGAGTTTCTTCTCAGGATCTTTTGTGCTATATTTCTGGCTAAAGAG 630
Db 361 GCTGCTGTAGAAAGGAGTTTCTTCTCAGGATCTTTTGTGCTATATTTCTGGCTAAAGAG 420
QY 631 AGAGTCTTTATGCCAGGACTCACTTTTCCAAATGAACTCATCAGCAGAGATGAAGACTT 690
Db 421 AGAGTCTTTATGCCAGGACTCACTTTTCCAAATGAACTCATCAGCAGAGATGAAGACTT 480
QY 691 CACTGTGACTTTGCTTGCCTGATGTTCCAAATGAACTCATCAGCAGAGATGAAGACTT 750
Db 481 CACTGTGACTTTGCTTGCCTGATGTTCCAAATGAACTCATCAGCAGAGATGAAGACTT 540
QY 751 GTCAAGGAGATCATTTGCTGCTGCAAAATTTGACAGAGTGTATTAACAGAGCTTG 810
Db 541 GTCAAGGAGATCATTTGCTGCTGCAAAATTTGACAGAGTGTATTAACAGAGCTTG 600
QY 811 CCAGTTGGCTCATTTGGAATGAATTTGATGAAACAGTACATTCAGTTTGTAGCT 870
Db 601 CCAGTTGGCTCATTTGGAATGAATTTGATGAAACAGTACATTCAGTTTGTAGCT 660
QY 871 GACAGATTAATTTGGAATGATGTTCTCAAAAGTTTTCAGCAGAGAAATCCTTTTGAT 930
Db 661 GACAGATTAATTTGGAATGATGTTCTCAAAAGTTTTCAGCAGAGAAATCCTTTTGAT 718
QY 931 TTTATGGAACATTT 946
Db 719 TATGGAACATTTCTTT 734
```

```
RESULT 14
BM801298
LOCUS BM801298
DEFINITION AGENCOURT_6422860 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5586693
5', mRNA sequence.
ACCESSION BM801298
VERSION BM801298.1 GI:19118121
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1061)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
```

cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12354 row: p column: 22
High quality sequence stop: 631.

FEATURES

source

1..1061
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5586693"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_92"
/note="Organ: testis; Vector: pCMV-SPORT6; Site: 1: NotI; Site 2: Sali; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 62.2%; Score 655.4; DB 4; Length 1061;
Best Local Similarity 97.6%; Pred. No. 2.4e-163;
Matches 718; Conservative 0; Mismatches 12; Indels 6; Gaps 5;

QY 1 ATGGGCGACCCGGAAAGCGCGGCTGGATCAGATGAGAGATCATCTTCA 60
Db 85 ATGGGCGACCCGGAAAGCGCGGCTGGATCAGATGAGAGATCATCTTCA 144
QY 61 GACACCAAGCAAGTCAAAATGAAGTCAAAATGAAGAGCACTCTCTAAGAAAGAGTTCTCGC 120
Db 145 GACACCAAGCAAGTCAAAATGAAGTCAAAATGAAGAGCACTCTCTAAGAAAGAGTTCTCGC 204
QY 121 CGGTTTGTCTATCTTTCCAAATCCAGTACCTCTGATTTTGGAAAATGTATAAACAGGCACAG 180
Db 205 CGGTTTGTCTATCTTTCCAAATCCAGTACCTCTGATTTTGGAAAATGTATAAACAGGCACAG 264
QY 181 GCTTCCTTCTGACAGCAGAGAGTCTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 240
Db 265 GCTTCCTTCTGACAGCAGAGAGTCTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 324
QY 241 CTTAAGCAGATCAGAGAGTACTTCACTCTCACATCTAGCCCTTTTTCAGCCAGTAT 300
Db 325 CTTAAGCAGATCAGAGAGTACTTCACTCTCACATCTAGCCCTTTTTCAGCCAGTAT 384
QY 301 CGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGAGTCCAGAGGCT 360
Db 385 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGAGTCCAGAGGCT 444
QY 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCAGAGATTTTCACTCAGAGATGTACAGTTTG 420
Db 445 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCAGAGATTTTCACTCAGAGATGTACAGTTTG 504
QY 421 CTGATAGACACTTACATCAGATCCCAAGAAAAGGAATTTTATTTATGCAATTTGAA 480
Db 505 CTGATAGACACTTACATCAGATCCCAAGAAAAGGAATTTTATTTAATGCAATTTGAA 564
QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGCCCTTGGATGGATAGCAGATAGAAAA 540
Db 565 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGCCCTTGGATGGATAGCAGATAGAAAA 624
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTTGTCTGTAGAAAGAGTGTCTTCTCTCAGGA 600
Db 625 TCTACTTTTGGGGAAGAGTGGTGGCTTTTGTCTGTAGAAAGAGTGTCTTCTCTCAGGA 684
QY 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGCTTTATGCGCAGACTCATCTTTTTC 660
Db 685 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGAGGCTC-TATGCGCAGGACTCACTTTTTC 743
QY 661 AATGAACATCAT-CAGCAGAGAT-GAAGGACTTCACTGTGA-CTTTGCTTGGCTGAT-GT 715

Db 744 CATGGACTATCCACAGAGATGGAGGACTTCCCTGTGAATTTGCCCTTGCTGTATGTT 803
QY 716 TCCAATACCTTAGTAAA 731
Db 804 TCCAATACCTTAGTAAA 819

RESULT 15
BUI70979
LOCUS BUI70979 896 bp mRNA linear EST 04-SEP-2002
DEFINITION AGENCOURT_7940338 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6143754
5', mRNA sequence.
ACCESSION BUI70979
VERSION BUI70979.1 GI:22684963
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 896)
NIH-MGC <http://mgi.nhl.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM13467 row: k column: 19
High quality sequence stop: 619.
Location/Qualifiers
1..896
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6143754"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 67"
/note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 1.75 kb. Library constructed by Life
Technologies."

ORIGIN
Query Match 62.2%; Score 654.8; DB 5; Length 896;
Best Local Similarity 98.7%; Pred. No. 3.4e-163;
Matches 671; Conservative 0; Mismatches 7; Indels 2; Gaps 1;
QY 1 ATGGGCGACCCGGAAGCCGGAAGCGCGCGCTGGATCAGATGAGAGATCATCTTCA 60
Db 55 ATGGGCGACCCGGAAGCCGGAAGCGCGCGCTGGATCAGATGAGAGATCATCTTCA 114
QY 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCCTAAGAAAGAGTTCTCGC 120
Db 115 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCCTAAGAAAGAGTTCTCGC 174
QY 121 CGTTTGTTCATCTTCCAAATCCAGTACCTGATATTTGGAAATGATATAAAGGACAG 180
Db 175 CGTTTGTTCATCTTCCAAATCCAGTACCTGATATTTGGAAATGATATAAAGGACAG 234
QY 181 GCTTCTCTTCGACAGCAGAGAGGTCCACTATCAAGAGTCTCCCTCACTGGACAG 240
Db 235 GCTTCTCTTCGACAGCAGAGAGGTCCACTATCAAGAGTCTCCCTCACTGGACAG 294
QY 241 CTTAAGCAGATGAGAAGTACTCTCATCTCTCATCTCTAGCTTTTGGCCAGCTGAT 300
Db 295 CTTAAGCAGATGAGAAGTACTCTCATCTCTCATCTCTAGCTTTTGGCCAGCTGAT 354

QY 301 GGAATGTAAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
Db 355 GGAATGTAAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 414
QY 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAATGTTCACTCAGAGATGTACAGTTTG 420
Db 415 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAATGTTCACTCAGAGATGTACAGTTTG 474
QY 421 CTGATAGACACTTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 475 CTGATAGACACTTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA 534
QY 481 ACCATGCCCTATGCTTAAGAAAAAGCAGATTGGGCCCTTGGATGGATAGCAGATAGAAAA 540
Db 535 ACCATGCCCTATGCTTAAGAAAAAGCAGATTGGGCCCTTGGATGGATAGCAGATAGAAAA 594
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCCCTTGGCTGTAGAAAGGAGTTTCTTCTCAGGA 600
Db 595 TCTACTTTTGGGGAAGAGTGGTGGCCCTTGGCTGTAGAAAGAGTTTCTTCTCAGGA 654
QY 601 TCTTTTGGCTGTATATTCTGGCTTAAAGAAAGAGAGGTCTTATGCCAGGACTCACATTTT 658
Db 655 TCTTTTGGCTGTATATTCTGGCTTAAAGAAAGAGAGGTCTTATGCCAGGACTCACATTTT 714
QY 659 CCAATGAACCTCATCAGCAGA 678
Db 715 CAATGAACCTCATCAGCAGA 734

Search completed: October 30, 2005, 03:23:14
Job time : 2726.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2005, 23:30:35 ; Search time 146 Seconds
(without alignments)
11801.366 Million cell updates/sec

Title: US-10-698-228-2

Perfect score: 1053

Sequence: 1 atggcgccaccggaaggcc.....tcacctggatgcagatttt 1053

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

- Issued Patents NA:*
- 1: /cgn2_6/ptodata/1/ina/5A-COMB.seq:*
 - 2: /cgn2_6/ptodata/1/ina/5B-COMB.seq:*
 - 3: /cgn2_6/ptodata/1/ina/6A-COMB.seq:*
 - 4: /cgn2_6/ptodata/1/ina/6B-COMB.seq:*
 - 5: /cgn2_6/ptodata/1/ina/PTCUS-COMB.seq:*
 - 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	1053	100.0	1053	4	US-10-019-733-2
2	1053	100.0	4955	4	US-10-019-733-3
3	1051.4	99.8	1053	4	US-10-019-733-12
4	1051.4	99.8	1081	4	US-10-019-733-4
5	579	55.0	2500	4	US-09-962-665-9
6	579	55.0	2500	4	US-09-963-333-9
7	579	55.0	2500	4	US-09-962-677-9
8	578.4	54.9	2479	4	US-09-949-016-2025
9	578.4	54.9	2500	4	US-09-023-655-1370
10	578.4	54.9	2500	4	US-09-949-016-145
11	441	41.9	481	3	US-08-905-223-125
12	429.8	40.8	14176	1	US-08-307-499-1
13	429.8	40.8	14176	1	US-08-307-499-14
14	429.8	40.8	14176	3	US-09-299-268-1
15	429.8	40.8	14176	3	US-09-299-268-14
16	383.2	36.4	1236	4	US-09-248-796A-3914
17	382.8	36.4	825	1	US-08-307-499-40
18	382.8	36.4	825	3	US-09-299-268-40
19	337.8	32.1	1112	1	US-08-136-743B-1
20	318.2	30.2	419	3	US-08-905-223-129
21	316	30.0	1083	4	US-09-248-796A-3915
22	260.8	24.8	656	3	US-09-328-111-791
23	195.2	18.5	449	4	US-09-270-767-14670
24	169.2	16.1	366	1	US-08-307-499-13
25	169.2	16.1	366	3	US-09-299-268-13
26	141.4	13.4	373	4	US-09-401-064-337
27	139.8	13.3	373	4	US-09-401-064-319

C	28	115.2	10.9	366	4	US-09-401-064-307	Sequence 307, Appl
	29	112.4	10.7	297	4	US-09-313-294A-4435	Sequence 4435, Ap
	30	93.8	8.9	11820	4	US-09-949-016-11887	Sequence 11887, A
	31	93.8	8.9	11826	4	US-09-949-016-13767	Sequence 13767, A
	32	92	8.7	601	4	US-09-949-016-20615	Sequence 20615, A
	33	92	8.7	601	4	US-09-949-016-69028	Sequence 69028, A
	34	74.4	7.1	276	4	US-09-313-294A-3437	Sequence 3437, Ap
	35	70.2	6.7	304	4	US-09-313-294A-7143	Sequence 7143, Ap
	36	68	6.5	276	4	US-09-313-294A-4638	Sequence 4638, Ap
	37	61	5.8	1053	4	US-09-902-540-2666	Sequence 2666, Ap
	38	61	5.8	13706	4	US-09-902-540-1124	Sequence 1124, Ap
	39	60	5.7	601	4	US-09-949-016-69031	Sequence 69031, A
	40	59.8	5.7	1141	4	US-09-806-708B-22	Sequence 22, Appl
	41	54.4	5.2	288	4	US-09-313-294A-6108	Sequence 6108, Ap
	42	52.2	5.0	276	4	US-09-313-294A-3244	Sequence 3244, Ap
	43	50	4.7	123025	4	US-09-198-452A-1	Sequence 1, Appli
	44	50	4.7	1230230	4	US-09-438-185A-1	Sequence 1, Appli
	45	45.4	4.3	124884	4	US-09-661-596A-76	Sequence 76, Appl

ALIGNMENTS

RESULT 1

US-10-019-733-2

; Sequence 2, Application US/10019733

; Patent No. 6682917

; GENERAL INFORMATION:

; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA

; TITLE OF INVENTION: New Protein and its DNA

; FILE REFERENCE: 2619WOOP

; CURRENT APPLICATION NUMBER: US/10/019,733

; CURRENT FILING DATE: 2001-12-28

; PRIOR APPLICATION NUMBER: JP 11-181131

; PRIOR FILING DATE: 1999-06-28

; PRIOR APPLICATION NUMBER: JP 11-192391

; PRIOR FILING DATE: 1999-07-06

; PRIOR APPLICATION NUMBER: JP 2000-017770

; PRIOR FILING DATE: 2000-01-21

; NUMBER OF SEQ ID NOS: 14

; SEQ ID NO 2

; LENGTH: 1053

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION:

US-10-019-733-2

Query Match

Best Local Similarity 100.0%; Score 1053; DB 4; Length 1053;

Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAGGCGGAGCGGCGGCGGCTGGATCAGGATGAGATCATCTTCA 60

Db 1 ATGGGCGACCCGGAAGGCGGAGCGGCGGCGGCTGGATCAGGATGAGATCATCTTCA 60

QY 61 GACACCAACGAAAGTGAATAAAGTCAAATGAAGAGCCACTCTTAAGAAAGTTCGCG 120

Db 61 GACACCAACGAAAGTGAATAAAGTCAAATGAAGAGCCACTCTTAAGAAAGTTCGCG 120

QY 121 CGGTTTGTCATCTTTCCTCAATCCAGTACCTGATATTTGGAATGTATTAACAGGCACAG 180

Db 121 CGGTTTGTCATCTTTCCTCAATCCAGTACCTGATATTTGGAATGTATTAACAGGCACAG 180

QY 181 GCTTCTCTTCTGACAGCAGAGAGGTCGACTTATCAAGATCTCCTCACTGGAAACAG 240

Db 181 GCTTCTCTTCTGACAGCAGAGAGGTCGACTTATCAAGATCTCCTCACTGGAAACAG 240

QY 241 CTTAAAGCAGATGAGAAGTACTTTCATCTCTCAATCTTAGCCCTTTTTCAGCCAGTGAT 300

Db 241 CTTAAAGCAGATGAGAAGTACTTTCATCTCTCAATCTTAGCCCTTTTTCAGCCAGTGAT 300

QY 301 GGAATTGTAATGAAANNTTGGTGGAGCCCTTTAGTCAGGAGTGCGAGGCTTCCAGGGCT 360

Db 301 GGAATTGTAATAAGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGTCAGAGGCT 360
QY 361 CGCTGTTTCTATGGCTTTCAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGGCTTTCAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 480
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATGCGCTTGGATGAGATAGAGATAGAAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAAAGCAGATGCGCTTGGATGAGATAGAGATAGAAAA 540
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAGAGGATTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAGAGGATTTCTTCTCAGGA 600
QY 601 TCTTTTGTCTATATTCTGGCTTAAAGAAAGAGAGGCTTTATGCCAGGACTCACCTTTTCC 660
Db 601 TCTTTTGTCTATATTCTGGCTTAAAGAAAGAGAGGCTTTATGCCAGGACTCACCTTTTCC 660
QY 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGGCTGTGTTCCAA 720
Db 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGGCTGTGTTCCAA 720
QY 721 TACTTAGTAATAAAGCCTTCAGAAAGAGGTCAGGAGATCAATGTTGATGCTGTCAAA 780
Db 721 TACTTAGTAATAAAGCCTTCAGAAAGAGGTCAGGAGATCAATGTTGATGCTGTCAAA 780
QY 781 ATTGACGAGGATTTTAAACAGAAAGCCTTGCAGTTGGCTCATTTGGAATGAATTCATT 840
Db 781 ATTGACGAGGATTTTAAACAGAAAGCCTTGCAGTTGGCTCATTTGGAATGAATTCATT 840
QY 841 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTAATTTGTGGAACCTTGGATTCTCA 900
Db 841 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTAATTTGTGGAACCTTGGATTCTCA 900
QY 901 AGGTTTTTTCAGCAGAAAAATCCTTTTGAATTTTATGAGAAAAACATTTCTTTAGAAGAAAA 960
Db 901 AGGTTTTTTCAGCAGAAAAATCCTTTTGAATTTTATGAGAAAAACATTTCTTTAGAAGAAAA 960
QY 961 ACAAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGGTTATGGCAGAAACC 1020
Db 961 ACAAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGGTTATGGCAGAAACC 1020

RESULT 2
US-10-019-733-3
; Sequence 3, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619W00P
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 3
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens

; FEATURE:
; OTHER INFORMATION:
US-10-019-733-3
Query Match 100.0%; Score 1053; DB 4; Length 4955;
Best Local Similarity 100.0%; Pred. No. 2,1e-310;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGCGACCCGAAAGCGCGGCGGCTGGATCAGATGAGAGATCATCTTCA 60
Db 245 ATGGGCGACCCGAAAGCGCGGCGGCTGGATCAGATGAGAGATCATCTTCA 304
QY 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCACTCTCTAAGAAAGAGTTCTCGC 120
Db 305 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCACTCTCTAAGAAAGAGTTCTCGC 364
QY 121 CGTTTGTCTATCTTCCAAATCCAGTACCTTGAATTTGGAAAAATGTATAAACAGGCACAG 180
Db 365 CGTTTGTCTATCTTCCAAATCCAGTACCTTGAATTTGGAAAAATGTATAAACAGGCACAG 424
QY 181 GCTTCCTTCTGGACAGCAGAGAGGTCGACTTATCAAGAGATCTCCCTCACTGGAACAAG 240
Db 425 GCTTCCTTCTGGACAGCAGAGAGGTCGACTTATCAAGAGATCTCCCTCACTGGAACAAG 484
QY 241 CTTAAAGCAGATCAGAAAGTACTTCTCATCTCTCAATCTTAGCCCTTTTGTGAGCCAGTGAT 300
Db 485 CTTAAAGCAGATCAGAAAGTACTTCTCATCTCTCAATCTTAGCCCTTTTGTGAGCCAGTGAT 544
QY 301 GGAATTTGTAATAAGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAAGTTCAGAGGCT 360
Db 545 GGAATTTGTAATAAGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAAGTTCAGAGGCT 604
QY 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAAGTTCACATCAGAGATGTACAGTTTG 420
Db 605 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAAGTTCACATCAGAGATGTACAGTTTG 664
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTCGAA 480
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTCGAA 724
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCCCTTGGATGAGATAGAGATAAAA 540
Db 725 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCCCTTGGATGAGATAGAGATAAAA 784
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAAAGAGGATTTTCTTCTCAGGA 600
Db 785 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAAAGAGGATTTTCTTCTCAGGA 844
QY 601 TCTTTTGTCTATATTCTGGCTTAAAGAAAGAGAGGTCCTATGCCAGGACTCACCTTTTCC 660
Db 845 TCTTTTGTCTATATTCTGGCTTAAAGAAAGAGAGGTCCTATGCCAGGACTCACCTTTTCC 904
QY 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTTGTGCTGCTGATGTTCCAA 720
Db 905 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTTGTGCTGCTGATGTTCCAA 964
QY 721 TACTTAGTAATAAAGCCTTCAGAAAGAGGTCAGGAGATCAATGTTGATGCTGTCAAA 780
Db 965 TACTTAGTAATAAAGCCTTCAGAAAGAGGTCAGGAGATCAATGTTGATGCTGTCAAA 1024
QY 781 ATTGACGAGGATTTTAAACAGAAAGCCTTGCAGTTGGCTCATTTGGAATGAATTCGATT 840
Db 1025 ATTGACGAGGATTTTAAACAGAAAGCCTTGCAGTTGGCTCATTTGGAATGAATTCGATT 1084
QY 841 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTAATTTGTGGAACCTTGGATTCTCA 900
Db 1085 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTAATTTGTGGAACCTTGGATTCTCA 1144
QY 901 AGGTTTTTTCAGCAGAAAAATCCTTTTGAATTTTATGAGAAAAACATTTCTTTAGAAGAAAA 960
Db 1145 AGGTTTTTTCAGCAGAAAAATCCTTTTGAATTTTATGAGAAAAACATTTCTTTAGAAGAAAA 1204
QY 961 ACAAAATTTCTTTGAGAAACAGGTTTCAGAGTATCAGCGTTTTCAGGTTATGGCAGAAACC 1020

Db 1205 ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1264
QY 1021 ACAGATAACGCTCTTACCTTTGGATGCGAGATTTT 1053
Db 1265 ACAGATAACGCTCTTACCTTTGGATGCGAGATTTT 1297

RESULT 3

US-10-019-733-12
; Sequence 12, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 12
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-12

Query Match 99.8%; Score 1051.4; DB 4; Length 1053;
Best Local Similarity 99.9%; Pred. No. 2.7e-310;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATGAGATCATCTTCA 60
Db 1 ATGGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATGAGATCATCTTCA 60
QY 61 GACACCAACGAAAGTGAATAAGTCAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAACGAAAGTGAATAAGTCAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
QY 121 CGGTTTGTCTTCCAAATCCAGTACCTGTATTTGAAAATGATATAACAGGCACAG 180
Db 121 CGGTTTGTCTTCCAAATCCAGTACCTGTATTTGAAAATGATATAACAGGCACAG 180
QY 181 GCTTCCTTCTGGACAGACAGAGGTGCGACTTATCAAGGATCTCCCTCACTGGAAACAAG 240
Db 181 GCTTCCTTCTGGACAGACAGAGGTGCGACTTATCAAGGATCTCCCTCACTGGAAACAAG 240
QY 241 CTTAAAGCAGATGAGAAAGTACTTCTCATCTCTAGACCTTTTTCAGCCAGTAT 300
Db 241 CTTAAAGCAGATGAGAAAGTACTTCTCATCTCTAGACCTTTTTCAGCCAGTAT 300
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTCCAGAGGCT 360
Db 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTCCAGAGGCT 360
QY 361 CGCTGTTTCTATGGCTTTTCAATTTCTCATCGAATGTTTCTCATGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGGCTTTTCAATTTCTCATCGAATGTTTCTCATGAGATGTACAGTTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATTGAA 480
QY 481 ACCATGCCCTATGTTTAAAGAAAAAGCAGATTTGGCCCTTGGATGATGATAGAGAAAA 540
Db 481 ACCATGCCCTATGTTTAAAGAAAAAGCAGATTTGGCCCTTGGATGATGATAGAGAAAA 540

QY 541 TCTACTTTTGGGAAAGAGTGGTGGCCTTTGCTGCTGTAGAAAGAGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCCTTTGCTGCTGTAGAAAGAGTTTCTTCTCAGGA 600
QY 601 TCTTTTGGCTGCTATATTTCTGGCTTAAAGAAAGAGAGTCTTATGCCAGAGTCACTTTTCC 660
Db 601 TCTTTTGGCTGCTATATTTCTGGCTTAAAGAAAGAGAGTCTTATGCCAGAGTCACTTTTCC 660
QY 661 AATGAACTCATCAGCAGAGATGAAGACTTTCACCTGTGACCTTTGCTGCTGATGTTCCAA 720
Db 661 AATGAACTCATCAGCAGAGATGAAGACTTTCACCTGTGACCTTTGCTGCTGATGTTCCAA 720
QY 721 TACTTAGTAATAAAGCCTTCAGAAAGAAAGGCTCAGGAGATCATTTGTGATGCTCTCAAA 780
Db 721 TACTTAGTAATAAAGCCTTCAGAAAGAAAGGCTCAGGAGATCATTTGTGATGCTCTCAAA 780
QY 781 ATTGAGCAGAGTCTTTTAAACAGAAAGCCTTCCAGTTGGCCTCATTTGGAATGAATTCAT 840
Db 781 ATTGAGCAGAGTCTTTTAAACAGAAAGCCTTCCAGTTGGCCTCATTTGGAATGAATTCAT 840
QY 841 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTAATTTGGAAGTGGATTTCTCA 900
Db 841 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTAATTTGGAAGTGGATTTCTCA 900
QY 901 AAGGTTTTTTCAGGCAGAAATCCTTTTGAATTTTATGAAAACATTTCTTTAGAAAGAAAA 960
Db 901 AAGGTTTTTTCAGGCAGAAATCCTTTTGAATTTTATGAAAACATTTCTTTAGAAAGAAAA 960
QY 961 ACAAAATTTCTTTGAGAAACAGAGTTTCAGAGTATCAGCGTTTTGCGATTATGGCAGAAACC 1020
Db 961 ACAAAATTTCTTTGAGAAACAGAGTTTCAGAGTATCAGCGTTTTGCGATTATGGCAGAAACC 1020
QY 1021 ACAGATAACGCTCTTCACTTGGATGCGAGATTTT 1053
Db 1021 ACAGATAACGCTCTTCACTTGGATGCGAGATTTT 1053

RESULT 4

US-10-019-733-4
; Sequence 4, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 4
; LENGTH: 1081
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-4

Query Match 99.8%; Score 1051.4; DB 4; Length 1081;
Best Local Similarity 99.9%; Pred. No. 2.7e-310;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATGAGATCATCTTCA 60
Db 20 ATGGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATGAGATCATCTTCA 79
QY 61 GACACCAACGAAAGTGAATAAGTCAAGAGCCACTCTTCAAGAAAGAGTTCTCGC 120
Db 80 GACACCAACGAAAGTGAATAAGTCAAGAGCCACTCTTCAAGAAAGAGTTCTCGC 139

```
QY 121 CGTTTGTTCATCTTCCCAATCCAGTACCTGATATTTGGAAATGTATAAACAGGCACAG 180
Db |||
QY 140 CGTTTGTTCATCTTCCCAATCCAGTACCTGATATTTGGAAATGTATAAACAGGCACAG 199
Db |||
QY 181 GCTTCTCTCTGACAGCAGAGAGGTGCGATTTATCAAGGATCTCCCTCACTGGAAACAAG 240
Db |||
QY 200 GCTTCTCTCTGACAGCAGAGAGGTGACTTATCAAGGATCTCCCTCACTGGAAACAAG 259
QY 241 CTTAAGCAGATGAGAGTACTCTCTCTCACTCTAGCCCTTTTTCAGCCAGTGTAT 300
Db |||
QY 260 CTTAAGCAGATGAGAGTACTCTCTCTCACTCTAGCCCTTTTTCAGCCAGTGTAT 319
QY 301 GGAATTTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGTTCAGAGGCT 360
Db |||
QY 320 GGAATTTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGTTCAGAGGCT 379
QY 361 CGCTGTTCTATGGCTTTCAAAATTCATCGAGAAATGTTCTACTCAGAGATGACAGTTTG 420
Db |||
QY 380 CGCTGTTCTATGGCTTTCAAAATTCATCGAGAAATGTTCTACTCAGAGATGACAGTTTG 439
QY 421 CTGATAGACACTATACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db |||
QY 440 CTGATAGACACTATACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 499
QY 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTCGGCTTCGCGATGGATAGCAGATAGAAAA 540
Db |||
QY 500 ACCATGCCCTATGTTAAGAAAAAAGCAGATTCGGCTTCGCGATGGATAGCAGATAGAAAA 559
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTGTCTGTAGAGAGGATTTTCTCTCAGGA 600
Db |||
QY 560 TCTACTTTTGGGAAAGAGTGGTGGCTTGTCTGTAGAGAGGATTTTCTCTCAGGA 619
QY 601 TCTTTTGTGCTATATTTCTGGCTTAAGAAAGAGAGTCTTATGCCAGGACTACTTTTCC 660
Db |||
QY 620 TCTTTTGTGCTATATTTCTGGCTTAAGAAAGAGAGTCTTATGCCAGGACTACTTTTCC 679
QY 661 AATGAACCTCATCAGCAGACATCAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db |||
QY 680 AATGAACCTCATCAGCAGACATCAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 739
QY 721 TACTTAGTAAATTAAGCTTTCAGAAGAAAGGTCAGGAGATCATTTGATGCTGTCAAA 780
Db |||
QY 740 TACTTAGTAAATTAAGCTTTCAGAAGAAAGGTCAGGAGATCATTTGATGCTGTCAAA 799
QY 781 ATTGAGCAGGAGTTTTAAACAGAGCCTTGCAGTTGGCTCATTTGGAATGAATTGCAAT 840
Db |||
QY 800 ATTGAGCAGGAGTTTTAAACAGAGCCTTGCAGTTGGCTCATTTGGAATGAATTGCAAT 859
QY 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTTACTTGTGGAATTCGATTTCTCA 900
Db |||
QY 860 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTTACTTGTGGAATTCGATTTCTCA 919
QY 901 AAGGTTTTTTCAGGAGAAAAATCCTTTGATTTTATGGAATAATTTCTTTAGAAAGAAAA 960
Db |||
QY 920 AAGGTTTTTTCAGGAGAAAAATCCTTTGATTTTATGGAATAATTTCTTTAGAAAGAAAA 979
QY 961 ACAAAATTTCTTTGAGAAACAGTTTCAGAGTATACAGCGTTTTCAGATTATGCGAGAAACC 1020
Db |||
QY 980 ACAAAATTTCTTTGAGAAACAGTTTCAGAGTATACAGCGTTTTCAGATTATGCGAGAAACC 1039
QY 1021 ACAGATAACGTTCTTCACTTGGATGAGATTTT 1053
Db |||
QY 1040 ACAGATAACGTTCTTCACTTGGATGAGATTTT 1072
```

RESULT 5

```
US-09-962-665-9
; Sequence 9, Application US/09962665
; Patent No. 6537759
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: POLYLPOLYGLUTAMATE SYNTHETASE GENE SEQUENCE
```

```
; TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING THE
; FILE REFERENCE: 11926-015004
; CURRENT APPLICATION NUMBER: US/09/962,665
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 128,1464
; OTHER INFORMATION: n = g or a
; NAME/KEY: misc_feature
; LOCATION: 189
; OTHER INFORMATION: n = t or g
; NAME/KEY: misc_feature
; LOCATION: 524
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc_feature
; LOCATION: 1399
; OTHER INFORMATION: n = t or a
; NAME/KEY: misc_feature
; LOCATION: 1636,1738,2259
; OTHER INFORMATION: n = c or t
; US-09-962-665-9
```

Query Match 55.0%; Score 579; DB 4; Length 2500;

Best Local Similarity 75.0%; Pred. No. 6.4e-166;

Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

```
QY 89 ATGAGAGCCACTCTCTAAGAAAGAGTTCTCGCGGTTTGTTCATCTTTCCAAATCCAGTACC 148
Db |||
QY 397 AGATGAGCGCTGCTGAGAGAAACCCCGCGCTTTGTTCATCTTCCCATCCAGTACC 456
QY 149 CTGATATTTGGAAAAATGTATAAACAGGCACAGGCTTCTCTGACAGCAGAGAGGTG 208
Db |||
QY 457 ATGATATCTGGCAGATGTATAGAAGCAGAGAGGCTTCTTTTGGACCGCCGAGAGGTG 516
QY 209 ACTTATCAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAAGTCTTCACT 268
Db |||
QY 517 ACCTCTCNAAGACATTTACAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTATAT 576
QY 269 CTCACATCTTAGCCTTTTTCAGCCAGTGAATTTGTAATGAAATTTTGGTGAGC 328
Db |||
QY 577 CCATGTTCTGGCTTTCTTTGCAAGGATGATGATGATGATGATGATGATGATGATGATGAT 636
QY 329 GCTTTAGTCAGAGGTGCGAGGTTCCAGAGGCTCGCTGTTTCTATGCTTTTCAAATTTCTCA 388
Db |||
QY 637 GATTTAGCCAAAGATTGAGATTACAGAAGCCGCTGTTTCTATGCTTCCAAATTTGCCA 696
QY 389 TCGAGAAATTTCACTCAGAGATGTAAGTTTGTGTATAGACACTTATCATCAGAGATCCCA 448
Db |||
QY 697 TGGAAAAACATACATTTCTGAAATGTATAGTCTTCTTATGACACTTACATAAAGATCCCA 756
QY 449 AGAAAAAGGAAATTTTATTTAATGCAATTTGAACCAATGCCCTATGTTAAGAAAAACAG 508
Db |||
QY 757 AAGAAAGGGAATTTTCTTTCAATGCAATTTGAACCAATGCCCTATGTTAAGAAAAACAG 816
QY 509 ATTGGGCTTTGCGATGATGAGAGATAGAAAAATCTACTTTTGGGAAAGAGTGTGGCCT 568
```

```
Db 817 ACTGGGCTTGGCTGGATGGGACAAAGAGGCTACCTATGTGTAACGTGTTGTAGCCT 876
Qy 569 TTGCTGCTAGAGGAGTTTCTCTCAGGATCTTTTGGCTGCTATATCTGGCTAAAGA 628
Db 877 TTGCTGCTAGAGGCAATTTCTTTCCGGTCTTTTGGTGGATATCTGGCTCAAGA 936
Qy 629 AGAGAGCTTATGCCAGAGCTACATTTTCCAAATCAATCAATCAATCAATCAATCAAT 688
Db 937 AACGAGGACTGATGCCCTGCCTCACATTTTCTAAATGAATTTATAGCAGATGAGGTT 996
Qy 689 TTCACCTGACCTTTGCTGCTGATGTTTCCAAATCAATCAATCAATCAATCAATCAAT 748
Db 997 TACACTGTGATTTGCTTGCCTGATGTTTCAAAACACCTGCTACAAACCATCGAGGAGA 1056
Qy 749 GGGTCAGGAGATCATTTGTTGATGCTGTCAAAATTCAGCAGGAGTTTAAACAGAGCCT 808
Db 1057 GAGTAAGAGAAATATATCAATGCTTTCGGATAGAACAGGAGTTCTCCTCAGGCTT 1116
Qy 809 TGCCAGTTGGCTCATTTGGAATGAATTTGATGAAACAGTACATTTGATGTTGTAG 868
Db 1117 TGCTGTGAAGCTCATTTGGATGAATTTGCACTTAATGAAGCAATATCATTGAGTTTGG 1176
Qy 869 CTGACAGATTACTTTGTGGAACCTTGGATTCTCAAAGGTTTTTTCAGGCAAGAAATCCTTTT 928
Db 1177 CAGACAGACTTATGCTGGAACCTGGGTTTTAGCAAGGTTTTTCAGAGTAGAGAACCCATTG 1236
Qy 929 ATTTATGAAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAACGAGTTTCAG 988
Db 1237 ACTTTATGAGAAATATTTCACTGGAAGGAAAGACTAACTTCTTTGAGAGAGTAGGCG 1296
Qy 989 AGTATCAGAGCTTTTGAGTATGGCAGAAACACAGATAACGTTCTTCACTTGGATGAG 1048
Db 1297 AGTATCAGAGGATGGAGTGAATGTCAGTCCAAACAGAGAAATCTTTACCTTGGATGCTG 1356
Qy 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 6
US-09-963-333-9
; Sequence 9, Application US/09963333
; Patent No. 6664062
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: THYMIDINE SYNTHASE GENE SEQUENCE VARIANCES
; TITLE OF INVENTION: HAVING UTILITY IN DETERMINING THE TREATMENT
; FILE OF INVENTION: OF DISEASE
; FILE REFERENCE: 11926-015002
; CURRENT APPLICATION NUMBER: US/09/963,333
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 128, 1464
; OTHER INFORMATION: n = g or a
; NAME/KEY: misc_feature
; LOCATION: 189
```

```
; OTHER INFORMATION: n = t or g
; NAME/KEY: misc_feature
; LOCATION: 524
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc_feature
; LOCATION: 1399
; OTHER INFORMATION: n = t or a
; NAME/KEY: misc_feature
; LOCATION: 1636-1738, 2259
; OTHER INFORMATION: n = c or t
US-09-963-333-9
```

Query Match 55.0%; Score 579; DB 4; Length 2500;

Best Local Similarity 75.0%; Pred. No. 6.4e-166;

Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

```
Qy 89 ATGAAGAGCCACTCTTAAGAAAGAGTTCTCGCGGTTTGTCACTTTTCCAAATCCAGTACC 148
Db 397 AGGATGAGCGCTGCTGAGAGAAAAACCCCGCGCTTGTTCATCTTCCCATCCGAGTACC 456
Qy 149 CTGATATTTGAAAATGTATAACAGGCACAGGCTTCTTCTGACACAGCAGAGGTCG 208
Db 457 ATGATATCTGGCAGATGTATAAGAGGCACAGGCTTCTTCTTGGACCGCCGAGGAGTTG 516
Qy 209 ACTTATCAAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAAGTACTTCACT 268
Db 517 ACCTCTCNAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT 576
Qy 269 CTCACATCTTAGCCTTTTGTGACCCAGTGAATTTGAAATGAAATTTGGTGGAGC 328
Db 577 CCCATGTTCTGGCTTCTTTTGCAGCAAGCGATGCGATGAAATGAAATTTGGTGGAGC 636
Qy 329 GCTTTAGTACGAGGTCAGGTTCCAGAGGCTCGCTGTTCTATGGCTTTCAAATTTCTCA 388
Db 637 GATTTAGCCAAAGATTCAGATTTACAGAGCCGCTGTTTCTATGGCTTCCAAATTTGCCA 696
Qy 389 TCGAGAAATGTTCACTCAGAGATGTACAGTTTGTGTAGACACTTTACATCAGAGATCCCA 448
Db 697 TGGAAACATACATCTGAAATGTATAGTCTTCTTATTGACACTTACATAAAGATCCCA 756
Qy 449 AGAAAAGGGAATTTTATTTAATGCAATTTGAAACCATGCTTATGTTTAAAGAAAAAGCAG 508
Db 757 AGAAAAGGGAATTTTCTTCTCAATGCCATGAAACGATGCTTGTGTCAAGNAGAGGAGC 816
Qy 509 ATTGGCCCTTGCATGAGTAGAGATAGAAAATCTACTTTTGGGAAAAGAGTGGTGCCCT 568
Db 817 ACTGGGCCCTTGCCTGGAATTTGGGACAAAGAGGCTACCTATGTTGTAACGTTGTAGCCT 876
Qy 569 TTGCTGCTGAGAGGAGTTTCTTCTCAGGATCTTTTGTGCTATATTTCTGGCTAAGA 628
Db 877 TTGCTGAGTGAAGGCAATTTTCTTTTCCGGTCTTCTTTGCGTCAATATCTCCTCAGGCT 936
Qy 629 AGAGAGGCTTATGCCAGGACTCACATTTTTCCTCAATGAACCTCATCAGCAGAGATCAAGGAC 688
Db 937 AACGAGAGCTGATGCTTGCCTCACATTTTCTAATGAATTTATAGCAGAGATGAGGTT 996
Qy 689 TTCACCTGATCTTTGCTTGCCTGATGTTTCCAAATCAATTTAGTAAATTAAGCCTTCAGAAGAAA 748
Db 997 TACACTGTGATTTTGTGCTGCTGATGTTTCAAAACACCTGGTACACAAACCATCGAGGAGA 1056
Qy 749 GGGTCAGGAGATCATTTGATGCTGTCAAAATTCAGCAGGAGTGTTCCTAACAGAGCCT 808
Db 1057 GAGTAAGAGAAATATATCAATGCTGTTTGGATAGAACAGGAGTTCTCCTCAGGCTT 1116
Qy 809 TGCCAGTTGGCTCATTTGGAATGAATTTGATGAAACAGTACATTTGAGTTTGTAG 868
Db 1117 TGCTGTGAAGCTCATTTGGATGAATTTGCACTTAATGAAGCAATATCATTGAGTTTGG 1176
Qy 869 CTGACAGATTACTTTGTGGAACCTTGGATTCTCAAAGGTTTTTTCAGGCAAGAAATCCTTTT 928
Db 1177 CAGACAGACTTATGCTGGAACCTGGGTTTTAGCAAGGTTTTTCAGAGTAGAGAACCCATTG 1236
Qy 929 ATTTATGAAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAACGAGTTTCAG 988
```


Db 1237 ACTTTATGAGAATATTTCACTGGAAGGAAAGACTAACTTCTTTGAGAAGAGTAGGCG 1296
QY 989 AGTATCAGCGTTTTCGAGTTATGGCAGAAACACACAGATAACGTCCTTCCACCTGGATGACG 1048
Db 1297 AGTATCAGAGGATGGAGTGATGTCAAGTCCACACAGAGAATTTCTTTACCTTTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 7

US-09-962-677-9
; Sequence 9, Application US/09962677
; Patent No. 6759200
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: THYMIDINE PHOSPHORYLASE GENE SEQUENCE
; TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING
; FILE OF INVENTION: THE TREATMENT OF DISEASE
; FILE REFERENCE: 11926-015003
; CURRENT APPLICATION NUMBER: US/09/962,677
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 128..1464
; OTHER INFORMATION: n = g or a
; NAME/KEY: misc_feature
; LOCATION: 189
; OTHER INFORMATION: n = t or g
; NAME/KEY: misc_feature
; LOCATION: 524
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc_feature
; LOCATION: 1399
; OTHER INFORMATION: n = t or a
; NAME/KEY: misc feature
; LOCATION: 1636..1738, 2259
; OTHER INFORMATION: n = c or t
US-09-962-677-9

Query Match 55.0%; Score 579; DB 4; Length 2500;
Best Local Similarity 75.0%; Pred. No. 6.4e-166;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;
QY 89 ATGAAGAGCCATCTTAAGAAGAGTTCTCGCGGTTTGATCTTTTCAATCCAGTACC 148
Db 397 AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTTGTCTATCTTCCCATCGAGTACC 456
QY 149 CTGATATTTGGAATGTATACAGGACAGGCTTCTTCTGGACAGAGAGGTCG 208
Db 457 ATGATATCTGGCAGATGTATAAGAGGAGAGGCTTCTTTTGGACCGCGGAGGTTG 516
QY 209 ACTTATCAAGAGTCTCCCTCACTGGAACAAGTTAAAGCAGATGAGAGTACTTCACT 268
Db 517 ACCTCTCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGATATTTATAT 576

QY 269 CTCACATCTTAGCCTTTTTCAGCCAGTCGATGGAATTTGTAATAAATTTGGTGGAGC 328
Db 577 CCCATGTTCTGGCTTTCTTTTGCAGCAAGCGATGCGATAGTAATAAATTTGGTGGAGC 636
QY 329 GCTTTAGTCAGGAGTGCAGGTTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAAATCTCA 388
Db 637 GATTTAGCCAAGAAGTTTCAGATTACAGAAGCCGCTGTTTCTATGGCTTCCAAATTTGCCA 696
QY 389 TCGAGAAATGTTCACTCAGAGATGTACAGTTTGTCTGTAGACACTTACATCAGAGATCCCA 448
Db 697 TCGAAAACATACATTTCTGAAATGTATAGTCTTCTTATTGACACTTACATAAAGATCCCA 756
QY 449 AGAAAAGGCAATTTTATTATTAATGCAATTTGAAACCATGCGCTATGTTTAAAGAAAAAGCAG 508
Db 757 AGAAAAGGCAATTTCTTCAATGCGCATTTGAACGATGCGCTTGTCTCAGAAAGAGGCGAG 816
QY 509 ATTGGGCCCTTGGCATGATAGCAGATAGAAAATCTACTTTTGGGAAAAGAGTGGTGGCCT 568
Db 817 ACTGGGCCCTTGGCTGGAATTTGGGCAACAAAGAGGCTACTATGTTGTAACGTTGTAGCCT 876
QY 569 TTGCTGCTGTAGAGGAGTTTCTTCTCAGGATCTTTTGTCTGTCTATATTTCTGGCTAAGA 628
Db 877 TTGCTGCACTGGAAGGCAATTTCTTTTCCGGTTCTTTTGGCTCGATATTTCTGGCTCAAGA 936
QY 629 AGAGAGGCTCTTATGCCAGGACTCACCTTTTCCAAATGAACCTCATCAGCAGAGATCAAGGAC 688
Db 937 AAGGAGGCTGATGCTGCGCTCACATTTTCTTAAGAACTTATTAGCAGAGATGAGGGTT 996
QY 689 TTCACCTGTGACTTTTGTCTGCTGATGTTTCAATCTTAGTAAATTAAGCCTTTCAGAGAAA 748
Db 997 TACACTGTGATTTTGTCTGCTGATGTTTCAAAACACCTGGTACACAAACCATCGGAGAGA 1056
QY 749 GGGTCAGGAGATCATTTGTTGATGCTGTCAAAATTTGAGCAGGAGGTTTTCACAGAGCCT 808
Db 1057 GAGTAAGAGAAAATAATTATCAATGCTGTTTGGGATAGAACAGGAGTTTCTCTCAGGCT 1116
QY 809 TGCCAGTTCGCTCATTTGGAATGAATTTGCAATTTTGTGAAACAGTACATGAGTTTGTAG 868
Db 1117 TGCTGTGAGCTCATTTGGATGAATTTGCACTCTTAATGAAGCAATATGAGTTTGTGG 1176
QY 869 CTGACAGATTACTTGTGGAATTTGGATTTCTCAAAGGTTTTCAGGCGAGAAAATCCTTTTG 928
Db 1177 CAGACAGACTTATGCTGGAATCTGGGTTTGTAGCAAGGTTTTCAGAGTAGAGAACCAATTG 1236
QY 929 ATTTATGGAATAACATTTCTTTAGAGGAAAACAAATTTCTTTGAGAAACGAGTTTCAG 988
Db 1237 ACTTTATGAGAATAATTTCACTGGAAGGAAAGACTAACTTCTTTGAGAAGAGAGTAGGCG 1296
QY 989 AGTATCAGCGTTTTCAGATTATGCGAGAAACACAGATAACGCTTCCACCTTGGATGCGAG 1048
Db 1297 AGTATCAGAGGATGGGAGTGTCAAGTCCAAACAGAGAAATTTCTTTTACCTTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 8

US-09-949-016-2025
; Sequence 2025, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03

```

; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 2025
; LENGTH: 2479
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-2025

```

Query Match	54.9%	Score	578.4	DB 4	Length	2479			
Best Local Similarity	75.0%	Pred. No.	9.6e-166						
Matches	723	Conservative	0	Mismatches	241	Indels	0	Gaps	0

QY	89	ATCAAGAGCCACTCC	TAAGAAAGAGTTCTCGCGGTTTGT	CATCTTTC	CAATCCAGTACC	148
Db	397	AGATGAGCGCGCTG	CTGAGAGAAACCCCGCGCTTTGT	CATCTTCC	CATCGAGTACC	456
QY	149	CTGATATTTGGAAAA	TGTATAAACAGGCACAGCGTCTCTT	CTGACAGCAGAGAGTGC	208	
Db	457	ATGATATCTGGCAG	TGTATAGAAAGCAGAGGCTTCTTTT	TGGACCGCGAGGAGTGG	516	
QY	209	ACTTATCAAAAGGA	TCCCTCACTCGAAACAAAGCTTAA	AGCAGATGAGAACTTCATCT	268	
Db	517	ACCTCTCCAAGGAC	ATTACGACATGGGAATCCCTGAAAC	CCGAGGAGAGATATTTTATAT	576	
QY	269	CTCACATCTTAGCC	TTTTTTTGGAGCCAGTGATGGAAATGT	AAATGAAAAATTTGGTGGAGC	328	
Db	577	CCCATGTTCTGGC	TTCTTTTGGACGAAGCATAGGCATAG	TAAATGAAAACTTGGTGGAGC	636	
QY	329	GCTTTTAACTCAGG	AGGTGCGAGTTTCCAGAGGCTCGCTGT	TTTCTATGGCTTTCAAAATCTCA	388	
Db	637	GATTTAGCCNAGAG	TTTCAGATTCAGAGACCCGCTGTTTCT	ATATGGCTTTCAAAATTCGCCA	696	
QY	389	TCGAGAAATGTTCT	CAGAGATGTA	CAGTTTGTCTGATAGACACTT	ATACAGAGATCCCA	448
Db	697	TGGAACAATACAT	TTCTGAAATGTATAGTCTTCTTAT	TGACACTTACATAAAAGATCCCA	756	
QY	449	AGAAAGGGAATTTT	TATTTTAAATGCAATTTGAAACCA	TGCCCTATGTTTAAAGAAAAAGCAG	508	
Db	757	AAGAAAGGGAATP	TTCTTCAATGGCAATTTGAAACGAT	TGCCCTGTGTCAAGAAAGAAAGCAG	816	
QY	509	ATTGGGCTTGGCAT	GGATAGCAGATAGAAAAATCTACT	TTTGGGGAAGAGTGTGTGCGCT	568	
Db	817	ACTGGGCTTGCCT	CGATTTGGGGAACAAGAGGCTAC	TATGTTGTAACGTGTTGTAGCCT	876	
QY	569	TTGCTGCTGTAGA	AGAGATTTTCTTCTCAGGATCTTTT	TGCTGTATATTTCTGGCTAAAGA	628	
Db	877	TTGCTCAGTGA	GAAGCAATTTCTTTTCGGTCTTTT	TGCGTCTGATATTTCTGGCTCAAGA	936	
QY	629	AGAGAGGCTTATG	CCAGGACTCAGTTTTTCCAAATGAA	CTCATCAGCAGAGATGAAGCAC	688	
Db	937	AACAGGAGCTAG	TGCTGGCCTCACATTTTCTAATGA	ACTTATTTAGCAGAGATGAGGGGT	996	
QY	689	TTCACTGTCACTT	TGCTGCTGATGTTCCAAATCTTAG	TAAATTAAGCCPTTCAGAAAGAA	748	
Db	997	TACACTGTGAT	TTTGTGCTGATGTTTCAACACCT	GTGTACAAACCATCGAGGAGA	1056	
QY	749	GGGTCAAGGAGAT	CAATTTGTGTGTGTGTGTGTGTGT	GTGTGTGTGTGTGTGTGTGTGTGTGT	808	
Db	1057	GAGTAAGAGAAAT	AAATTAATCAATGCTGTTCCGATAG	AAACAGGAGTTTCTCACTGAGGCT	1116	
QY	809	TGCCAGTTGGCCT	CATTTGGAATTTGCAATTTTGTAT	GAAACAGTACATTTGAGTTGTAG	868	
Db	1117	TGCTGTGAAAGT	CAATTTGGGATGAAATTTGCACTCT	TAATGAAGCAATATTTGAGTTGTGTG	1176	
QY	869	CTGACAGATTA	CTGTGGAACTTGGAAATCTCAAA	AGGTTTTTTCAGGCAAGAAAAATCT	928	
Db	1177	CAGACAGACTTA	TGCTGGAAGCTGGGTTTTAGC	AGGTTTTTTCAGAGTAGAACCCATTG	1236	
QY	929	ATTTTATGGA	AAACATTTCTTTAGAAAGGAAAA	CAAAATTTCTTTTGTGAAACGAGTTTCAG	988	
Db	1237	ACTTTATGAGA	ATAATTTTCACTTGGAGGAAAGACT	TAATCTTTTGTGAAAGAGAGTAGGCG	1296	

Qy	989	AGTATCAGCGTTTTTGCGATTATGGCAGAAACCA	CACAGATAACGTC	TTTCACTTGGATGCAG	1048
Db	1297	AGTATCAGAGGATGGGAGTGATGTC	AAAGTCCACAGAGAA	TTCTTTTACCTTGGATGCTG	1356
Qy	1049	ATTT	1052		
Db	1357	ACTT	1360		

RESULT 9

```

US-09-023-655-1370
; Sequence 1370, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Sellhammer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSER: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1370:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2500 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g36154
US-09-023-655-1370

```

Query Match	54.9%;	Score 578.4;	DB 4;	Length 2500;
Best Local Similarity	75.0%;	Pred. No. 9.7e-166;		
Matches 723;	Conservative	0;	Mismatches 241;	Indels 0; Gaps 0;
QY	89	ATGAAGAGCCACTCCTAAAGAAAGAGTTCTCGCGGTTTTGTCACTTTTCCAAATCCAGTACC	148	
Db	397	AGGATGACCGCTGCTGAGAGAAACCCCGCGCTTTGTCACTTCCCCATCGAGTACC	456	
QY	149	CTGATATTTGGAAATGTATAACAGGCACAGGCTTCCTTCTGGACAGCAGAAGAGGTCG	208	
Db	457	ATGATATCTGGCAGATGTATAGAGGCAGAGGCTTCTCTTTTGGACCGCCGAGGAGGTTG	516	
QY	209	ACTTATCAAGGATCTCCCTCACTGGAAACAAGCTTTAAAGCAGATGAGAAGTACTTTCATCT	268	

Db 517 ACCTCTCCAGGACATTCAGCATCTGGAAATCCCTGAAACCCGAGGAGAGATATTTTATAT 576
QY 269 CTCACATCTTAGCTTTTTCAGCCAGTGTGAATTTGTAATGAAATTTGGTGGAGC 328
Db 577 CCATGTTCTGGCTTCTTTGACGACGATGGCATAGTAATGAAATTTGGTGGAGC 636
QY 329 GCTTTAGTCAGAGGTGCGAGTTCCAGAGGCTGGCTGTTTCTATGCTTTTCAAAATTTCTCA 388
Db 637 GATTTAGCCAAGAAAGTTTCAATACAGAGGCGCTGTTTCTATGCTTTCCAAATTTGCCA 696
QY 389 TCGAGAAATGTTCACTCAGAGATGTACAGTTTGTGTATAGACATTCATCAGAGATCCCA 448
Db 697 TCGAAATACATATCTGAAATGTATAGTCTTTTATGACATTCATCAATAAAGATCCCA 756
QY 449 AGAAAGGCAATTTTATTTAATGCAATTTGAACCATGCGCTATGTTAAGAAAGAGAG 508
Db 757 AGAAAGGCAATTTTCTTCAATGCAATTTGAACCATGCGCTATGTTAAGAAAGAGAG 816
QY 509 ATTGGGCTTGGCATGATAGCAGATAGAAATCTACTTTTGGGGAAGAGTGGTGGCT 568
Db 817 ACTGGGCTTGGCATGATAGCAGATAGAAATCTACTTTTGGGGAAGAGTGGTGGCT 876
QY 569 TTGCTGCTAGAGAGTTTCTTCTAGAGTCTTTTCTAGAGTCTTTTCTAGAGTCTTTTCTAG 628
Db 877 TTGCTGCTAGAGAGTTTCTTCTAGAGTCTTTTCTAGAGTCTTTTCTAGAGTCTTTTCTAG 936
QY 629 AGAGAGTCTTATGCCAGGACTCACATTTTCCAAATGCAATCTATCAGCAGAGATGAAAGG 688
Db 937 AAGAGGACTGATGCTGCGCTCACATTTCTTAATGAATTTATGAGAGATGAGGGT 996
QY 689 TTCACATGTGACTTTGCTGCTGATGTTCCAAATCTTAGTAATAAGCCTTTCAGAAAGAA 748
Db 997 TACACTGTGATTTTGTGCTGATGTTTCAACACCTGTTTACAAACCATCGAGGAG 1056
QY 749 GGTTCAGGAGATCATTTGTTGATGCTGCTCAAAATTTGATGAAACAGTACATGAGTTGTAG 808
Db 1057 GAGTAAGAGAAATTAATATCAATGCTGTTCCGGATAGAACAGGAGTTCTCTAGGAGC 1116
QY 809 TGCCAGTGGCCTCATTTGGAATGAAATTTGATGAAACAGTACATGAGTTGTAG 868
Db 1117 TGCCAGTGGCCTCATTTGGAATGAAATTTGATGAAACAGTACATGAGTTGTAG 1176
QY 869 CTGACAGATTTCTGTTGGAATTTGATTTCTCAAGGTTTTCAGGAGAAATCTTTTGG 928
Db 1177 CTGACAGATTTCTGTTGGAATTTGATTTCTCAAGGTTTTCAGGAGAAATCTTTTGG 1236
QY 929 ATTTATGGAATCATTTTCTTTAGAGGAAATCAAAATTTCTTTGAGAAACGAGTTTTCAG 988
Db 1297 AGTATCAGAGGATGGGAGTGTCAAGTCCACAGAGAAATCTTTTACCTTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 10

US-09-949-016-145
; Sequence 145, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 145
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-145

Query Match 54.9%; Score 578.4; DB 4; Length 2500;
Best Local Similarity 75.0%; Pred. No. 9.7e-166;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

QY 89 ATGAAGAGCCACTCTAAGAAAGAGTTCTGCGGGTTTGTCTCATCTTCCAAATCCAGTACC 148
Db 397 AGATAGAGCGCTGCTGAGAGAAACCCCGCGCTTGTCTCATCTTCCCAATCCAGTACC 456
QY 149 CTGATATTTGGAAATGTATAACAGGCACAGGCTTCTTCTGACAGCAGAGAGAGTGC 208
Db 457 ATGATATCTGGCAGATGTATAGAGGCGAGAGGCTTCTTGGACCGCCGAGAGGTTG 516
QY 209 ACTTATCAAAAGATCTCCCTCACTGGAAACAAGCTTAAAGCAGATGAGAACTTCTCATCT 268
Db 517 ACCTCTCCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTATAT 576
QY 269 CTCACATCTTAGCTTTTTCAGCAGCAGTGTGAATTTGTAATGAAATTTGGTGGAGC 328
Db 577 CCATGTTCTGGCTTTTTCAGCAAGGATGATGATGATGATGATGATGATGATGATGATG 636
QY 329 GCTTTAGTCAGAGGTCAGAGTTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAATTTCTCA 388
Db 637 GATTTAGCCAGAGTTTCAAGATTCAGAGGCGCTGTTTCTATGGCTTCCAAATTTGCCA 696
QY 389 TCGAGAAATGTTCACTCAGAGATGTACAGTTTGTGTATAGACATTCATCAGAGATCCCA 448
Db 697 TCGAAATACATATCTGAAATGTATAGTCTTTCTTATGACACTTACATAAAGATCCCA 756
QY 449 AGAAAGGCAATTTTATTTAATGCAATTTGAACCATGCGCTATGTTAAGAAAGAGAG 508
Db 757 AAGAAAGGCAATTTTCTTCAATGCGCTTTGAAACGATGCTTGTGTCAAGAAAGAGAG 816
QY 509 ATTGGGCTTGGCATGATAGCAGATAGAAATCTACTTTTGGGGAAGAGTGGTGGCT 568
Db 817 ACTGGGCTTGGCATGATAGCAGATAGAAATCTACTTTTGGGGAAGAGTGGTGGCT 876
QY 569 TTGCTGCTAGAGAGTTTCTTCTCAGAGTCTTTTGTGCTATATTTCTGCTAAAGA 628
Db 877 TTGCTGCTAGAGAGTTTCTTCTCAGAGTCTTTTGTGCTATATTTCTGCTCAAGA 936
QY 629 AGAGAGTCTTATGCCAGGACTCACATTTTCCAAATGCAATCTATCAGCAGAGATGAAAGG 688
Db 937 AAGAGGACTGATGCTGCGCTCACATTTCTTAATGAATTTATGAGAGATGAGGGT 996
QY 689 TTCACATGTGACTTTGCTGCTGATGTTCCAAATCTTAGTAATAAGCCTTTCAGAAAGAA 748
Db 997 TACACTGTGATTTTGTGCTGATGTTTCAACACCTGTTTACAAACCATCGAGGAG 1056
QY 749 GGTTCAGGAGATCATTTGTTGATGCTGCTCAAAATTTGATGAAACAGTACATGAGTTGTAG 808
Db 1057 GAGTAAGAGAAATTAATATCAATGCTGTTCCGGATAGAACAGGAGTTCTCTAGGAGC 1116
QY 809 TGCCAGTGGCCTCATTTGGAATGAAATTTGATGAAACAGTACATGAGTTGTAG 868
Db 1117 TGCCAGTGGCCTCATTTGGAATGAAATTTGATGAAACAGTACATGAGTTGTAG 1176
QY 869 CTGACAGATTTCTGTTGGAATTTGATTTCTCAAGGTTTTCAGGAGAAATCTTTTGG 928
Db 1177 CTGACAGATTTCTGTTGGAATTTGATTTCTCAAGGTTTTCAGGAGAAATCTTTTGG 1236
QY 929 ATTTATGGAATCATTTTCTTTAGAGGAAATCAAAATTTCTTTGAGAAACGAGTTTTCAG 988
Db 1297 AGTATCAGAGGATGGGAGTGTCAAGTCCACAGAGAAATCTTTTACCTTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360

Db 1237 ACTTATGAGAAATATTTCACTGGAGGAAAGACTAACTTCTTTTGAGAGAGAGTAGGCG 1296
QY 989 AGTATCAGCGTTTTCAGTATATGGCAGAAACACAGATAACGCTTTCACCTTTGGATGCAG 1048
Db 1297 AGTATCAGAGGATGGAGTGAATGTCAGTCCAAACAGAGAAATCTTTTACCTTTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 11
US-08-905-223-125
; Sequence 125, Application US/08905223
; Patent No. 622029
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste D.
; APPLICANT: Duclert, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS
; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/905,223
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 125:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 481 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: DOUBLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; TISSUE TYPE: Brain
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 41..343
; IDENTIFICATION METHOD: Von Heijne matrix
; OTHER INFORMATION: score 4.4
; OTHER INFORMATION: seq ISHILAPFAASDG/IV
US-08-905-223-125

Query Match 41.9%; Score 441; DB 3; Length 481;
Best Local Similarity 100.0%; Pred. No. 3.2e-124;
Matches 441; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATGAGAGATCATCTTCA 60
Db 41 ATGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATGAGAGATCATCTTCA 100
QY 61 GACACCAACGAAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 101 GACACCAACGAAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 160

QY 121 CGTTTGTGTCATCTTTTCAATCCAGTACCTGATATTTGAAAAATGTATAAACAGGCACAG 180
Db 161 CGTTTGTGTCATCTTTTCAATCCAGTACCTGATATTTGAAAAATGTATAAACAGGCACAG 220
QY 181 GCTTCCTTTTGGACAGCAGAGAGGTGCACTTATCAAAAGGATCTCCCTCACTGGAACAAG 240
Db 221 GCTTCCTTTTGGACAGCAGAGAGGTGCACTTATCAAAAGGATCTCCCTCACTGGAACAAG 280
QY 241 CTTAAAGCAGATGAGAGTACTTCTCATCTCTCACATCTTAGCCCTTTTTCAGGCCAGTGAT 300
Db 281 CTTAAAGCAGATGAGAGTACTTCTCATCTCTCACATCTTAGCCCTTTTTCAGGCCAGTGAT 340
QY 301 GGAATTGTAATAATGAAAAATTTGGTGGAGCGCTTTTAGTCAGAGGTTGCAGGTTCAGAGGCT 360
Db 341 GGAATTGTAATAATGAAAAATTTGGTGGAGCGCTTTTAGTCAGAGGTTGCAGGTTCAGAGGCT 400
QY 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db 401 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 460
QY 421 CTGATAGACACTTACATCAGA 441
Db 461 CTGATAGACACTTACATCAGA 481

RESULT 12
US-08-307-499-1/c
; Sequence 1, Application US/08307499
; Patent No. 5651972
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; TITLE OF INVENTION: Live Vaccine Vector
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: U.S.A.
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/307,499
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: US 07/908,241
; FILING DATE: 1-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,630
; FILING DATE: 29-JUN-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/342,212
; FILING DATE: 21-APR-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF35.1.FWCC1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:

TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 138..1460
FEATURE:
NAME/KEY: CDS
LOCATION: 2456..2659
FEATURE:
NAME/KEY: CDS
LOCATION: 2809..3030
FEATURE:
NAME/KEY: CDS
LOCATION: 3070..3330
FEATURE:
NAME/KEY: CDS
LOCATION: 3356..4180
FEATURE:
NAME/KEY: CDS
LOCATION: 4392..5894
FEATURE:
NAME/KEY: CDS
LOCATION: 6171..6398
FEATURE:
NAME/KEY: CDS
LOCATION: 6447..6875
FEATURE:
NAME/KEY: CDS
LOCATION: 6928..7431
FEATURE:
NAME/KEY: CDS
LOCATION: 7454..7858
FEATURE:
NAME/KEY: CDS
LOCATION: 7895..8155
FEATURE:
NAME/KEY: CDS
LOCATION: 8215..8682
FEATURE:
NAME/KEY: CDS
LOCATION: 8715..9539
FEATURE:
NAME/KEY: CDS
LOCATION: 9562..10272
FEATURE:
NAME/KEY: CDS
LOCATION: 10316..11908
FEATURE:
NAME/KEY: CDS
LOCATION: 11971..12780
FEATURE:
NAME/KEY: CDS
LOCATION: 12829..13107
FEATURE:
NAME/KEY: CDS
LOCATION: 13149..14171
US-08-307-499-14

Query Match 40.8%; Score 429.8; DB 1; Length 14176;
Best Local Similarity 66.2%; Pred. No. 5.9e-120;
Matches 651; Conservative 0; Mismatches 327; Indels 5; Gaps 2;

QY 71 AAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCCCGGTTGTCA 130
|||||
Db 3336 AAATGATTTTATACAGAAATGGAGCCTATTCTTCAAGAGTCTGATCTAGGTTGTTA 3395
|||||
QY 131 TCTTTCCAAATCCAGTACCTGATATTGGAATATATAAAGCAGGCTTCTCTTCT 190
|||||
Db 3396 TTTTCCCTATTAAAGTATCATGATATCTGAAATATATAAACAATCAGTGGCAAGTTT 3455
|||||
QY 191 GGACAGCAGAGAGGTCGACTTATCAAGGATCTCCCTCACTGGAAACAAGCTTAAGCAG 250
|||||

Db 3456 GGACCGTTGAAGAGTAGATTTATCAAAAGATTAGATGATTGGGATAAAATTAACATAAG 3515
|||
QY 251 ATGAGAAGTACTTTCATCTCTCAGATCTTAGCCTTTTTCAGCCAGTATGGAATTGTAA 310
|||||
Db 3516 ACGAAAAATACCTTTTATAAAAAATATACTAGCATTTTTCAGATCTAGTATGTTAA 3575
|||||
QY 311 ATGAAAAATTTGGTGGAGCGCTTTAGTCAGAGGTCAGAGTTCCAGAGGCTCGCTGTTTCT 370
|||||
Db 3576 ATGAGAATTTAGCGGAAGATTTTATGTGATGTACAGTGTTCAGAGGCAGATGTTTCT 3635
|||||
QY 371 ATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTGTCTGATAGACA 430
|||||
Db 3636 ATGGATTTCAAAATAGCTATGAAAAATATTCTATTAGAAAAATGATAGTTTATTAATAGATA 3695
|||||
QY 431 CTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTTAATGCAATTTGAAACCATGCCCT 490
|||||
Db 3696 CATATGTAAGAGATAAATATAGAAAAAATGCAATTTTATTTAAACGCTATAGAAAAATGGAAT 3755
|||||
QY 491 ATGTTAAGAAAAAGCAGATTTGGGCTTGGGATGATAGAGATAGAGAAATCTACTTTTG 550
|||||
Db 3756 GCGTAAAAAAGAAAGCTGATTGGGCCAGAAAAATGGATATC---TAGCAACAAGGTATATG 3812
|||||
QY 551 GGGAAAGAGTGGTGGCTTTTGCTGTGTAGAAAGAGATTTTCTCTCAGGATCTTTTGTCTG 610
|||||
Db 3813 GAGAAAGATGATGATGATTTTGGAGCTGTGGAGGAATATTCTTTTCTGTTTCAATTTGCTG 3872
|||||
QY 611 CTATATTTCTGGCTAAAGAGAGAGGTTCTTATGCGAGACTCACATTTTCCAAATGAACTCA 670
|||||
Db 3873 CTATATTTTGGATAAAAAAACAAGGATTTGATGCGCGATTTAAACATTTTCTTAATGAACTAA 3932
|||||
QY 671 TCAGCAGAGATGAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCATACATTAGTAA 730
|||||
Db 3933 TAAGTAGAGACGAAGGTTTACATTTGTCATTTTTCGCTGTTTAAATGTTTAAACATTTATTATC 3992
|||||
QY 731 ATAAGCCTTCAGAAAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAAAATTCAGCAGG 790
|||||
Db 3993 ATCCACCATCTAAGGAAGTTATTAACGTCGATATCATTTGATCGGTTTATATAGAAAGG 4052
|||||
QY 791 AGTTTTTAAACAGAGCCTTGCAGTTGGCCTCATTTGGAATGAAATGCAATTTGATGAAAC 850
|||||
Db 4053 AGTTTTTGACAGTTGCTATTTCGCTGATCTTATAGTATGAAATTTGTTTAAATGCTC 4112
|||||
QY 851 AGTACATTGAGTTTGTAGCTGACAGATTTCTTGTGGAATTTGGAATTTTCAAGGTTTTC 910
|||||
Db 4113 AGTATATAGAAATTCGTCGAGATAGATTTATTAACAGAGTTAGGTTTGTGAAAG--TCTCA 4170
|||||
QY 911 AGCGAAAAATCCTTTTGTATTTTATGAAAAACATTTCTTTAGAGGAAAAACAATTTCT 970
|||||
Db 4171 ATGTATATATCTTTTGTAGCTTTTATGAGTATATATCCTAGAGGTTAGACTAAATTTT 4230
|||||
QY 971 TTGAGAAACGAGTTTCAGAGATATCAGCGTTTTCAGTTTTCAGTTATGCGAGAAACACAGATAACG 1030
|||||
Db 4231 TCGAACGACGAGTTAGTGAATATCAAAAGATGGGGTGTGTTTACAAATAAAGAGAGATA 4290
|||||
QY 1031 TCTTCACTTGGATGAGATTTT 1053
|||||
Db 4291 TATTTTCTACGGATATAGATTTT 4313
|||||

RESULT 14

US-09-299-268-1/c
; Sequence 1, Application US/09299268
; Patent No. 6217882
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, B.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; TITLE OF INVENTION: Live Vaccine Vector
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik

STREET: 2421 N.W. 41st Street, Suite A-1
 CITY: Gainesville
 STATE: Florida
 COUNTRY: U.S.A.
 ZIP: 32606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/299,268
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/901,127
 FILING DATE:
 APPLICATION NUMBER: US 07/908,241
 FILING DATE: 1-JUL-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/908,630
 FILING DATE: 29-JUN-1992
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/342,212
 FILING DATE: 21-APR-1992
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Salliwanchik, David R.
 REGISTRATION NUMBER: 31,794
 REFERENCE/DOCKET NUMBER: UF35.1.FWCC1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 904-375-8100
 TELEFAX: 904-372-5800
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 14176 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: unknown
 MOLECULE TYPE: DNA (genomic)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 3852..4226
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 4585..4887
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 5131..5310
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 5760..5912
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 6786..7130
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 10148..10513
 US-09-299-268-1

Query Match 40.8%; Score 429.8; DB 3; Length 14176;
 Best Local Similarity 66.2%; Pred. No. 5.9e-120;
 Matches 651; Conservative 0; Mismatches 327; Indels 5; Gaps 2;

QY 71 AAGTGAATAAGTCAATGAAGAGCCACTCTAAGAAAGAGTCTCGCCGGTTGTCA 130
 DB 10841 AAGTGAATTTTATACAGCAATGGAGCCCTATTCTTCAAGAGTCTGATTTAGTTCGTTA 10782
 QY 131 TCTTTCCATCCAGTACCTCGATATTGGAATATGTAATAACAGGCACAGGCTTCCTCT 190
 DB 10781 TTTTCCCTATTAAAGTATCATGATATCTGGAAAAATGTAATAACAATCAGTGGCAAGTTT 10722

QY 191 GGACAGCAGAAAGAGTGCAGCTTATCAAGAGATCTCCCTCACTCGAAACAAGCTTAAAGCAG 250
 DB 10721 GGACCGTTCGAAGAGTAGATTTATCAAAAGATTTAGATGATGGGATAAATTAACATAAG 10662
 QY 251 ATGAGAAGTACTTCTCATCTCTAGCCTTTTATGAGCCTTTTTCAGCAGCAGTGAATGTAA 310
 DB 10661 ACGAAAAATACCTTTATAAAACATATACATAGCATTTTTCGATCTAGTGTGATTTGTA 10602
 QY 311 ATGAAAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCTCGCTGTTCT 370
 DB 10601 ATGAGAATTTAGCGGAAGATTTATGTGATGTACAGTGTTCAGAGCAGATGTTTCT 10542
 QY 371 ATGGCTTTCAAAATCTCATCGAATGTTCACTCAGAGATGTACAGTGTTCGTGATAGACA 430
 DB 10541 ATGGATTTCAAAATAGCTATGGAATAATTCATTCAGAAATGTATAGTTTATTAATAGA 10482
 QY 431 CTTACATCAGATCCCAAGAAAGGAATTTTATTAATGCAATTTGAAACCAATGCGCT 490
 DB 10481 CATATGTAAGAGATAATATAGAAAAAATGCAATTTATTAACGCTATAGAAAAACAATGA 10422
 QY 491 ATGTTAAGAAAAAAGCAGATTTGGCCTTGCATGAGTAGCAGATAGAAAAATCTACTTTG 550
 DB 10421 GCGTAAAAAAGAAAGCTGATTTGGCCAGAAAAATGGATAIC---TAGCAACAAGTATATG 10365
 QY 551 GGGAAAGAGTGGTGGCCTTTTCTGCTGTAGAAAGAGTTTTCTTCTCAGAGATCTTTTCTG 610
 DB 10364 GAGAAAGAGTAGTAGCATTTGACGCTGTGGAGGGAATTTCTTTCTGGTTCAATTTGCTG 10305
 QY 611 CTATATCTGGCTAAAGAGAGAGGCTTTATGCGCAGGACTCACTTTTCCAAATGAACCTCA 670
 DB 10304 CTATATTTTGGATAAAAAAGCAGGATTTGATGCGCGGATTAACATTTTCTAATGAAC 10245
 QY 671 TCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGAGTGTCCCAATCTACTAGT 730
 DB 10244 TAAGTAGAGACGAAGGTTTACATTTGATTTGCGTGTGTTAATGTTTAAACATTTTAC 10185
 QY 731 ATAAGCCTTCAGAAAGAAAGGTCAGGAGATCAATTTGATGCTGTCAAAATTCAGCAGG 790
 DB 10184 ATCCACCATTAAGGAAGTTATAACGTCGATATCATTTGATGCGGTTAATATAGAAAAG 10125
 QY 791 AGTTTTTAACAGAAAGCCTTGGCCTTCAATGCGGCTCAATGGAATGAAATTTGATGAAC 850
 DB 10124 AGTTTTTACAGTGTGCTATTCCGGTGGATCTTATAGTATGAATTTGTTTAAATGCTC 10065
 QY 851 AGTACATTCAGTTTGTAGCTGCAGAGATTTCTGTGNACTTCTCAAGGTTTTC 910
 DB 10064 AGTATATGAATTCGTCGAGATAGATTTAACAAGAGTTAGGTTGTGAAAG--TCTCA 10007
 QY 911 AGGCAGAAATCCTTTTGTGATTTTATGAAAAACATTTCTTTAGAAAGAAAAACAAATTTCT 970
 DB 10006 ATGTATATATCTTTTGTAGCTTTTATGAGTATATATCATAGAAAGTTAGACTTAATTTT 9947
 QY 971 TTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTATTTGCGAGAAACCAACAGATAACG 1030
 DB 9946 TCGAACGAGAGTTAGTGAATATCAAAAGAGTGGGGTGTTCACAAATAAAGAGAGAATA 9887
 QY 1031 TCTTCACTTGGATGCAGATTTT 1053
 DB 9886 TATTTTCTACGGATATAGATTTT 9864

RESULT 15

US-09-299-268-14

; Sequence 14, Application US/09299268

; Patent No. 6217882

; GENERAL INFORMATION:

; APPLICANT: Moyer, Richard W.

; APPLICANT: Vi uela, Eladio

; APPLICANT: Gibbs, E.P.J.

; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a

; TITLE OF INVENTION: Live Vaccine Vector

; NUMBER OF SEQUENCES: 60

; CORRESPONDENCE ADDRESS:

ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: U.S.A.
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/299,268
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/901,127
FILING DATE:
APPLICATION NUMBER: US 07/908,241
FILING DATE: 1-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/908,630
FILING DATE: 29-JUN-1992
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/342,212
FILING DATE: 21-APR-1992
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UF35.1.FWCC1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 14176 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 138..1460
FEATURE:
NAME/KEY: CDS
LOCATION: 2456..2659
FEATURE:
NAME/KEY: CDS
LOCATION: 2809..3030
FEATURE:
NAME/KEY: CDS
LOCATION: 3070..3330
FEATURE:
NAME/KEY: CDS
LOCATION: 3356..4180
FEATURE:
NAME/KEY: CDS
LOCATION: 4392..5894
FEATURE:
NAME/KEY: CDS
LOCATION: 6171..6398
FEATURE:
NAME/KEY: CDS
LOCATION: 6447..6875
FEATURE:
NAME/KEY: CDS
LOCATION: 6928..7431
FEATURE:
NAME/KEY: CDS
LOCATION: 7454..7858
FEATURE:

NAME/KEY: CDS
LOCATION: 7895..8155
FEATURE:
NAME/KEY: CDS
LOCATION: 8215..8682
FEATURE:
NAME/KEY: CDS
LOCATION: 8715..9539
FEATURE:
NAME/KEY: CDS
LOCATION: 9562..10272
FEATURE:
NAME/KEY: CDS
LOCATION: 10316..11908
FEATURE:
NAME/KEY: CDS
LOCATION: 11971..12780
FEATURE:
NAME/KEY: CDS
LOCATION: 12829..13107
FEATURE:
NAME/KEY: CDS
LOCATION: 13149..14171
US-09-299-268-14

Query Match 40.8%; Score 429.8; DB 3; Length 14176;
Best Local Similarity 66.2%; Pred. No. 5.9e-120;
Matches 651; Conservative 0; Mismatches 327; Indels 5; Gaps 2;

QY 71 AAAGTGAATAAAGTCAAAATGAAGACCCACTCTCTAAGAAAGAGTCTCGCGGGTTTGTCA 130
|||
Db 3336 AAAATGATTTTATACACGAATGGAGCCCTATTCTTCAAGAGTCTGATTCTAGGTTTCGTTA 3395

QY 131 TCTTTCCAAATCCAGTACCTCGATATTTGGAAATGTATAAACAGGACACAGCTTCCTTCT 190
|||
Db 3396 TTTTCCCTATTAAAGTATCATGATATCTGGAAATGTATAAACATCAGTGCACAGTTT 3455

QY 191 GGACAGCAGAGAGGTCTGACTTATCAAGAGTCTCCCTCACTCGAACAAAGCTTAAAGCAG 250
|||
Db 3456 GGACCGTTGAAGAGTAGATTTATCAAAAGATTTAGATGTTGGGTAATAATCACTAAAG 3515

QY 251 ATGAGAAGTACTTCTCATCTCTCATCTTAGCCCTTTTTCAGGCAGTGATGGAATCTGAA 310
|||
Db 3516 ACGAAATATCTTTTAAACACATATCTAGCATTTTTCGATCTAGTGAATGTTGTTAA 3575

QY 311 ATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCAGAGGCTCGCTGTTTCT 370
|||
Db 3576 ATGAGAAATTTAGCGGAAAGATTTTATGTGGATGTACAGTGTTCAGAGGCACGATGTTCT 3635

QY 371 ATGGCTTTCAAAATCTCATCGAGAAATGTTTCACTCAGAGATGTACAGTTTCTGATAGACA 430
|||
Db 3636 ATGGATTTCAAAATAGCTATGGAATAATTTTCAATTCAGAAATGTATAGTTTAAATAGATA 3695

QY 431 CTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAAACCATGCGCT 490
|||
Db 3696 CATATGTAGAGATATATAGAAAAATGCAATTTATTAACGCTATAGAAACATGGAAT 3755

QY 491 ATGTTAAGAAAAAGCAGATTGCGCTTCGCGATGGATAGCAGATAGAAAAATCTACTTTTG 550
|||
Db 3756 GCGTAAAAAAGAAAGCTGATTGGCCAGAAAAATGGATATC---TAGCAACAGGATATG 3812

QY 551 GGGAAAGAGTGTGGCTTTGCTGTGTAGAGAGAGTTTCTTCTCAGGATCTTTTCTG 610
|||
Db 3813 GAGAAAGAGTAGTAGCATTTGCAGCTGTGGAGGGAATTTCTTTTCTGTTTCAATTTGCTG 3872

QY 611 CTATATCTCGCTAAAGAGAGAGGTCTTTATGCGAGACTCACTTTTCCAAATGAACCTCA 670
|||
Db 3873 CTATATTTTGGATAAAAAACGAGGATTTGATGCCCGGATTAACATTTTCTTAATGAACATA 3932

QY 671 TCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTTCCCAATACTAGTAA 730
|||
Db 3933 TAAGTAGAGACGAAGGTTTACATTTGATTTTGGCGTGTAAATGTTTAAACATTTATTAC 3992

QY	731	ATAAGCCTTCAGAGAAAGGTCAGGAGATCAATTGTTGATCTGTCAAAATTGAGCAGG	790
Db	3993	ATCCACCATCTAAGGAAGTTATAACGTCGATAATCATTTGATCGGTTAATATAGAAAAGG	4052
QY	791	AGTTTTTAACAGAACCTTGCAGTTGGCCTCATTTGGAATGAATTGCAATTTTGATGAAC	850
Db	4053	AGTTTTTGACAGTTGCTATTCGGTGGATCTTATAGGTATGAATTGTTGTTTAAATGCTC	4112
QY	851	AGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACTTGGATTCTCAAAAGTTTTTC	910
Db	4113	AGTATATAGAAATTCGTCAGATAGATTATTAAACAGATTAGTTGTGAAAAG--TCTCA	4170
QY	911	AGGCAGAAAATCCTTTTTCATTTTATGGAACAATTCTTTAGAGGAAAAACAAATTTCT	970
Db	4171	ATGTATATAATCCTTTTAGCTTTATGGAGTATATACACTAGAAGTAAGACTAATTTT	4230
QY	971	TTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGAGTTATGGCAGAAACCAAGATAAG	1030
Db	4231	TCGAAACGACGAGTTAGTGAATATCAAAAGATGGGGTGTTTACAAATAAAGAGAGAATA	4290
QY	1031	TCTTCACCTTGGATGCAGATTTT	1053
Db	4291	TATTTTCTACGGATATAGATTTT	4313

Search completed: October 30, 2005, 03:28:12
Job time : 150 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 30, 2005, 01:51:37 ; Search time 676.5 Seconds
(without alignments)
12855.457 Million cell updates/sec

Title: US-10-698-228-2

Perfect score: 1053

Sequence: 1 atggggcagccgaaagcc.....teacttgatgcagatttt 1053

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 9784742 seqs, 4129495052 residues

Total number of hits satisfying chosen parameters: 19569484

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Lasting first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq2:*
- 14: /cgn2_6/ptodata/1/pubna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubna/US10D_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/1/pubna/US10E_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/1/pubna/US10F_PUBCOMB.seq:*
- 20: /cgn2_6/ptodata/1/pubna/US10G_PUBCOMB.seq:*
- 21: /cgn2_6/ptodata/1/pubna/US10H_PUBCOMB.seq:*
- 22: /cgn2_6/ptodata/1/pubna/US10I_PUBCOMB.seq:*
- 23: /cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq:*
- 24: /cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq2:*
- 25: /cgn2_6/ptodata/1/pubna/US11A_PUBCOMB.seq:*
- 26: /cgn2_6/ptodata/1/pubna/US11_NEW_PUB.seq:*
- 27: /cgn2_6/ptodata/1/pubna/US16_NEW_PUB.seq:*
- 28: /cgn2_6/ptodata/1/pubna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1053	100.0	1053	19	US-10-698-228-2	Sequence 2, Appli
2	1053	100.0	4955	18	US-10-172-118-71	Sequence 71, Appl
3	1053	100.0	4955	19	US-10-342-887-71	Sequence 71, Appl
4	1053	100.0	4955	19	US-10-698-228-3	Sequence 3, Appli
5	1051.4	99.8	1053	19	US-10-698-228-12	Sequence 12, Appl

6	1051.4	99.8	1081	19	US-10-698-228-4	Sequence 4, Appli
7	578.4	54.9	1989	9	US-09-925-301-505	Sequence 505, App
8	578.4	54.9	2216	16	US-10-084-817-342	Sequence 342, App
9	578.4	54.9	2482	24	US-10-220-335-514	Sequence 514, App
10	578.4	54.9	2500	9	US-09-954-456-724	Sequence 724, App
11	578.4	54.9	2500	9	US-09-954-456-1169	Sequence 1169, Ap
12	578.4	54.9	2500	9	US-09-954-456-1827	Sequence 1827, Ap
13	578.4	54.9	2500	19	US-10-641-643-1370	Sequence 1370, Ap
14	578.4	54.9	2500	21	US-10-733-878-458	Sequence 458, App
15	578.4	54.9	2500	22	US-10-843-641A-3751	Sequence 3751, Ap
16	578.4	54.9	2500	22	US-10-843-641A-4196	Sequence 4196, Ap
17	578.4	54.9	2500	22	US-10-843-641A-4854	Sequence 4854, Ap
18	578.4	54.9	2500	24	US-10-756-149-713	Sequence 713, App
19	570.4	54.2	2113	22	US-10-764-420-1636	Sequence 1636, Ap
20	569	54.0	2641	24	US-10-220-335-170	Sequence 170, App
21	561	53.3	1328	19	US-10-403-571-75	Sequence 75, Appl
22	545.4	51.8	977	18	US-10-264-237-790	Sequence 790, App
23	514	48.8	1371	24	US-10-450-763-15278	Sequence 15278, A
24	463	44.0	1289	26	US-11-097-143-19181	Sequence 19181, A
25	447.8	42.5	186854	22	US-10-872-156-34	Sequence 34, Appl
26	442	42.0	1146	9	US-09-822-830A-174	Sequence 174, App
27	432.2	41.0	1218	16	US-10-128-714-7245	Sequence 7245, Ap
28	420.4	39.9	3945	26	US-11-097-143-19180	Sequence 19180, A
29	403.2	38.3	963	18	US-10-369-493-34052	Sequence 34052, A
30	397	37.7	1146	16	US-10-128-714-2245	Sequence 2245, Ap
31	392.6	37.3	1173	18	US-10-369-493-46252	Sequence 46252, A
32	392	37.2	1242	17	US-10-032-585-6537	Sequence 6537, Ap
33	386.6	36.7	1292	18	US-10-369-493-27809	Sequence 27809, A
34	380.6	36.1	1248	19	US-10-424-599-126843	Sequence 126843,
35	380.4	36.1	1200	18	US-10-369-493-45857	Sequence 45857, A
36	371.2	35.3	1361	19	US-10-424-599-40471	Sequence 40471, A
37	365.8	34.7	1586	19	US-10-424-599-73024	Sequence 73024, A
38	364.8	34.6	1206	18	US-10-320-797-2145	Sequence 2145, Ap
39	361.6	34.3	957	21	US-10-650-047-43	Sequence 43, Appl
40	351.4	33.4	629	24	US-10-450-763-6851	Sequence 6851, Ap
41	349	33.1	1314	16	US-10-128-714-1245	Sequence 1245, Ap
42	349	33.1	1450	16	US-10-128-714-6245	Sequence 6245, Ap
43	349	33.1	3314	16	US-10-128-714-245	Sequence 245, App
44	349	33.1	3450	16	US-10-128-714-5245	Sequence 5245, App
45	328.2	31.2	560	9	US-09-864-761-12381	Sequence 12381, A

ALIGNMENTS

RESULT 1

US-10-698-228-2
; Sequence 2, Application US/10698228
; Publication No. US2004007253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619W00P
; CURRENT APPLICATION NUMBER: US/10/698,228
; PRIOR FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-2

Query Match 100.0%; Score 1053; DB 19; Length 1053;

Db 1021 ACAGATAACGCTCTTCACCTTGGATGCAGATTTT 1053

RESULT 2

US-10-172-118-71

; Sequence 71, Application US/10172118

; Publication No. US20030224374A1

; GENERAL INFORMATION:

; APPLICANT: Dai, Hongyue

; APPLICANT: He, Yudong

; APPLICANT: Linsley, Peter

; APPLICANT: Mao, Mao

; APPLICANT: Roberts, Chris

; APPLICANT: Van 't Veer, Laura

; APPLICANT: Van de Vijver, Marc

; APPLICANT: Bernards, Rene

; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients

; FILE REFERENCE: 9301-175-999

; CURRENT APPLICATION NUMBER: US/10/172,118

; CURRENT FILING DATE: 2002-06-14

; PRIOR APPLICATION NUMBER: 60/380,770

; PRIOR FILING DATE: 2002-05-14

; NUMBER OF SEQ ID NOS: 2699

; SEQ ID NO 71

; LENGTH: 4955

; TYPE: DNA

; ORGANISM: Homo sapiens

; PUBLICATION INFORMATION:

; DATABASE ACCESSION NUMBER: AB036063

; DATABASE ENTRY DATE: 2001-06-18

US-10-172-118-71

Query Match 100.0%; Score 1053; DB 18; Length 4955;

Best Local Similarity 100.0%; Pred. No. 2.7e-294;

Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCGACCCGGAAGCGCGCGGCTGGATCAGATGAGAGATCATCTTCA 60

Db 245 ATGGCGACCCGGAAGCGCGCGGCTGGATCAGATGAGAGATCATCTTCA 304

Qy 61 GACACCAACGAAAGTGAATAAGTCAATGAAGGCCACTCCTAAGAAGAGTCTCGC 120

Db 305 GACACCAACGAAAGTGAATAAGTCAATGAAGGCCACTCCTAAGAAGAGTCTCGC 364

Qy 121 CGGTTTGTCTATCTTTCCAAATCCAGTACCCTGATATTTGGAAATGTATATAAAGAGTCTCGC 364

Db 365 CGGTTTGTCTATCTTTCCAAATCCAGTACCCTGATATTTGGAAATGTATATAAAGAGTCTCGC 424

Qy 181 GCTTCCTCTTGACAGCAGAGAGGTCGATCTTATCAAGGATCTCCCTCACTGGAAACAAG 240

Db 425 GCTTCCTCTTGACAGCAGAGAGGTCGATCTTATCAAGGATCTCCCTCACTGGAAACAAG 484

Qy 241 CTTAAAGCAGATGAGAGTACTTCTATCTCTCATCTTAGCCCTTTTTCGAGCCAGTGAT 300

Db 485 CTTAAAGCAGATGAGAGTACTTCTATCTCTCATCTTAGCCCTTTTTCGAGCCAGTGAT 544

Qy 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360

Db 545 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 604

Qy 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAATGTTTCACTCAGAGATGTACAGTTTG 420

Db 605 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAATGTTTCACTCAGAGATGTACAGTTTG 664

Qy 421 CTGATGACACTTACATCAGATGCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480

Db 665 CTGATGACACTTACATCAGATGCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 724

Qy 481 ACCATGCCCTATGTTTAAGAAAAAAGCAGATTGGGCTTCGCGATGGATAGCAGATAGAAAA 540

Db 725 ACCATGCCCTATGTTTAAGAAAAAAGCAGATTGGGCTTCGCGATGGATAGCAGATAGAAAA 784

Qy 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGTGTAGAGGAGTCTTCTCTCAGGA 600

Db 1021 ACAGATAACGCTCTTCACCTTGGATGCAGATTTT 1053

Db 785 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGTGCTGTAGAAAGGAGTCTTCTCTCAGGA 844
Qy 601 TCTTTTGTCTGTATATTCTGGCTTAAAGAGAGAGGTCTTATGCTCCAGGACTCCTCTTTTCC 660
Db 845 TCTTTTGTCTGTATATTCTGGCTTAAAGAGAGAGGTCTTATGCTCCAGGACTCCTCTTTTCC 904
Qy 661 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTCTTGCCTGTATGTTCCAA 720
Db 905 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTCTTGCCTGTATGTTCCAA 964
Qy 721 TACTTAGTAAATAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGTGATGCTGTCAA 780
Db 965 TACTTAGTAAATAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGTGATGCTGTCAA 1024
Qy 781 ATTGAGCAGGAGTCTTAAACAGAGCCTTGCAGTGGCTCATTTGGAATGAATTCATTT 840
Db 1025 ATTGAGCAGGAGTCTTAAACAGAGCCTTGCAGTGGCTCATTTGGAATGAATTCATTT 1084
Qy 841 TTGATGAACAGTACATTTAGTGTGTAGCTGACAGATTACTTGTGAACTTGGATTCTCA 900
Db 1085 TTGATGAACAGTACATTTAGTGTGTAGCTGACAGATTACTTGTGAACTTGGATTCTCA 1144
Qy 901 AAGGTTTTTCAGGAGAGAAATCCTTTTGATTTTATGGAAGAAATTTCTTTAGAGGAAA 960
Db 1145 AAGGTTTTTCAGGAGAGAAATCCTTTTGATTTTATGGAAGAAATTTCTTTAGAGGAAA 1204
Qy 961 ACAATTTCTTTGAGAACAGTTCAGAGTATCAGCGTTTTCAGTATGCGAGTATGCGAGAAAC 1020
Db 1205 ACAATTTCTTTGAGAACAGTTCAGAGTATCAGCGTTTTCAGTATGCGAGTATGCGAGAAAC 1264
Qy 1021 ACAGATAACGCTCTTACCTTGGATGCGAGATTTT 1053
Db 1265 ACAGATAACGCTCTTACCTTGGATGCGAGATTTT 1297

RESULT 3

US-10-342-887-71
; Sequence 71, Application US/10342887
; Publication NO. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 71
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-71

Query Match 100.0%; Score 1053; DB 19; Length 4955;
Best Local Similarity 100.0%; Pred. No. 2.7e-294;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCGCAGCCCGAAGCGCGGCGGCTGGATCAGATGAGAGATCATCTTCA 60
Db 245 ATGGCGCAGCCCGAAGCGCGGCGGCTGGATCAGATGAGAGATCATCTTCA 304

Qy 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTCTCTGC 120
Db 305 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTCTCTGC 364
Qy 121 CGTTTGTCTATCTTTTCCAAATCCAGTACCTGTATATTTGGAAATATGTATAACAGGACAG 180
Db 365 CGTTTGTCTATCTTTTCCAAATCCAGTACCTGTATATTTGGAAATATGTATAACAGGACAG 424
Qy 181 GCTTCTCTTGGACAGCAGAGAGAGGTGCACTTATCAAGAGATCTCCCTCACTGGAACAG 240
Db 425 GCTTCTCTTGGACAGCAGAGAGAGGTGCACTTATCAAGAGATCTCCCTCACTGGAACAG 484
Qy 241 CTTAAAGCAGATCAGAAAGTACTTCTCTCACTCTTAGCCTTTTGTGAGCCAGTGTAT 300
Db 485 CTTAAAGCAGATCAGAAAGTACTTCTCTCACTCTTAGCCTTTTGTGAGCCAGTGTAT 544
Qy 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGAGTGCAGAGGCT 360
Db 545 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGAGTGCAGAGGCT 604
Qy 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGATGTTTCACTCAGAGATGTACAGTTG 420
Db 605 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGATGTTTCACTCAGAGATGTACAGTTG 664
Qy 421 CTGATAGACACTTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 665 CTGATAGACACTTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 724
Qy 481 ACCATGCCCTATGTTAAGAAAGAGAGATTTGGGCTTGGATGGATAGCAGATAGAAA 540
Db 725 ACCATGCCCTATGTTAAGAAAGAGAGATTTGGGCTTGGATGGATAGCAGATAGAAA 784
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTAGAGAGATTTTCTTCTCAGGA 600
Db 785 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTAGAGAGATTTTCTTCTCAGGA 844
Qy 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGTCTTATGCCAGACTCACCTTTTCC 660
Db 845 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGTCTTATGCCAGACTCACCTTTTCC 904
Qy 661 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTCTGCTGTATGTTCCAA 720
Db 905 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTCTGCTGTATGTTCCAA 964
Qy 721 TACTTAGTAAATAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGTGATGCTGTCAA 780
Db 965 TACTTAGTAAATAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGTGATGCTGTCAA 1024
Qy 781 ATTGAGCAGGAGTCTTAAACAGAGCCTTGCAGTGGCTCATTTGGAATGAATTCATTT 840
Db 1025 ATTGAGCAGGAGTCTTAAACAGAGCCTTGCAGTGGCTCATTTGGAATGAATTCATTT 1084
Qy 841 TTGATGAACAGTACATTTAGTGTGTAGCTGACAGATTACTTGTGAACTTGGATTCTCA 900
Db 1085 TTGATGAACAGTACATTTAGTGTGTAGCTGACAGATTACTTGTGAACTTGGATTCTCA 1144
Qy 901 AAGGTTTTTCAGGAGAGAAATCCTTTTGATTTTATGGAAGAAATTTCTTTAGAGGAAA 960
Db 1145 AAGGTTTTTCAGGAGAGAAATCCTTTTGATTTTATGGAAGAAATTTCTTTAGAGGAAA 1204
Qy 961 ACAATTTCTTTGAGAACAGTTCAGAGTATCAGCGTTTTCAGTATGCGAGTATGCGAGAAAC 1020
Db 1205 ACAATTTCTTTGAGAACAGTTCAGAGTATCAGCGTTTTCAGTATGCGAGTATGCGAGAAAC 1264
Qy 1021 ACAGATAACGCTCTTACCTTGGATGCGAGATTTT 1053
Db 1265 ACAGATAACGCTCTTACCTTGGATGCGAGATTTT 1297

RESULT 4

US-10-698-228-3
; Sequence 3, Application US/10698228

```
; Publication No. US20040072253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 3
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-3

Query Match      100.0%; Score 1053; DB 19; Length 4955;
Best Local Similarity 100.0%; Pred. No. 2.7e-294;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
DB 245 ATGGGCGACCCGGAAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
QY 61 GACACCCAGAAAGTGAATTAAGTCAATTAAGAGCCACTCTTAAGAAGAGTTCCTCC 120
DB 305 GACACCAAGAAAGTGAATTAAGTCAATTAAGAGCCACTCTTAAGAAGAGTTCCTCC 364
QY 121 CGGTTTGTCATCTTCCATCCAGTACCTCTGATATTTGGAAATGTATAAACAGGCACAG 180
DB 365 CGGTTTGTCATCTTCCATCCAGTACCTCTGATATTTGGAAATGTATAAACAGGCACAG 424
QY 181 GCTTCCTCTTGACAGCAGAAAGTGCATCTATCAAGAGTCTCCCTCACTGGAACAAG 240
DB 425 GCTTCCTCTTGACAGCAGAAAGTGCATCTATCAAGAGTCTCCCTCACTGGAACAAG 484
QY 241 CTTAAAGCAGATGAGAGTACTTCTCTCATCTTACCTTTTTCGAGGCAGTGTAT 300
DB 485 CTTAAAGCAGATGAGAGTACTTCTCTCATCTTACCTTTTTCGAGGCAGTGTAT 544
QY 301 CGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGTGCAGGTTCCAGAGGCT 360
DB 545 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGTGCAGGTTCCAGAGGCT 604
QY 361 CGCTGTTTCTATGGCTTCCAAATTTCTCATCGAGATGTTTCACTCAGAGATGTACAGTTG 420
DB 605 CGCTGTTTCTATGGCTTCCAAATTTCTCATCGAGATGTTTCACTCAGAGATGTACAGTTG 664
QY 421 CTGATAGACACTATACAGATCCCAAGAAAGGAAATTTTATTTATGCAATTTGAA 480
DB 665 CTGATAGACACTATACAGATCCCAAGAAAGGAAATTTTATTTATGCAATTTGAA 724
QY 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGCCCTTCGATGGATAGCAGATAGAAA 540
DB 725 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGCCCTTCGATGGATAGCAGATAGAAA 784
QY 541 TCTACTTTTGGGGAAGAGTGTGGCTTTTGCTGTGTAGAGAGTGTTCCTCTCAGGA 600
DB 785 TCTACTTTTGGGGAAGAGTGTGGCTTTTGCTGTGTAGAGAGTGTTCCTCTCAGGA 844
QY 601 TCTTTTGTGCTATATCTGGCTTAAAGAGAGAGGCTTTATGCGCAGACTCACTTTTCC 660
DB 845 TCTTTTGTGCTATATCTGGCTTAAAGAGAGAGGCTTTATGCGCAGACTCACTTTTCC 904
QY 661 AATGAACATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGTATGTTCCA 920
; Publication No. US20040072253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 3
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-2

Query Match      99.8%; Score 1051.4; DB 19; Length 1053;
Best Local Similarity 99.8%; Pred. No. 3.1e-294;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
DB 1 ATGGGCGACCCGGAAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
QY 61 GACACCAAGAAAGTGAATTAAGTCAATTAAGAGCCACTCTTAAGAAGAGTTCCTCC 120
DB 61 GACACCAAGAAAGTGAATTAAGTCAATTAAGAGCCACTCTTAAGAAGAGTTCCTCC 120
QY 121 CGGTTTGTCATCTTCCATCCAGTACCTCTGATATTTGGAAATGTATAAACAGGCACAG 180
DB 121 CGGTTTGTCATCTTCCATCCAGTACCTCTGATATTTGGAAATGTATAAACAGGCACAG 180
QY 181 GCTTCCTCTTGAGCAGCAGAAAGAGTGCATCTTCAAGAGGATCTCCCTCACTGGAACAAG 240
; Publication No. US20040072253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 12
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-12

Query Match      99.8%; Score 1051.4; DB 19; Length 1053;
Best Local Similarity 99.8%; Pred. No. 3.1e-294;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
DB 1 ATGGGCGACCCGGAAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
QY 61 GACACCAAGAAAGTGAATTAAGTCAATTAAGAGCCACTCTTAAGAAGAGTTCCTCC 120
DB 61 GACACCAAGAAAGTGAATTAAGTCAATTAAGAGCCACTCTTAAGAAGAGTTCCTCC 120
QY 121 CGGTTTGTCATCTTCCATCCAGTACCTCTGATATTTGGAAATGTATAAACAGGCACAG 180
DB 121 CGGTTTGTCATCTTCCATCCAGTACCTCTGATATTTGGAAATGTATAAACAGGCACAG 180
QY 181 GCTTCCTCTTGAGCAGCAGAAAGAGTGCATCTTCAAGAGGATCTCCCTCACTGGAACAAG 240
```

```
Db 181 GCTTCCTCTGACAGCAGAGAGGTTGATCTATCAAGAGATCTCCCTCACTGGAAACAAG 240
Qy 241 CTTAAAGCAGATGAGAGTACTTCTCTCATCTCTACATCTTAGCCCTTTTTCAGCCAGTGAT 300
Db 241 CTTAAAGCAGATGAGAGTACTTCTCTCATCTCTACATCTTAGCCCTTTTTCAGCCAGTGAT 300
Qy 301 GGAAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
Db 301 GGAAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
Qy 361 CCGTGTCTATGCGCTTCAAAATCTCATCGAGAAATGTTTCACTCAGAGATGATACAGTTTG 420
Db 361 CCGTGTCTATGCGCTTCAAAATCTCATCGAGAAATGTTTCACTCAGAGATGATACAGTTTG 420
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 480
Qy 481 ACCATGCCCTATGTTAAAGAAAGAGAGATTGGGCTTCGATGGATAGATAGAGAA 540
Db 481 ACCATGCCCTATGTTAAAGAAAGAGAGATTGGGCTTCGATGGATAGATAGAGAA 540
Qy 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTGTGTAAGAGGATTTTCTTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTGTGTAAGAGGATTTTCTTCAGGA 600
Qy 601 TCTTTTGTGCTATATCTGCTTAAGAGAGAGGCTTATGCCAGGACTCACTTTTCC 660
Db 601 TCTTTTGTGCTATATCTGCTTAAGAGAGAGGCTTATGCCAGGACTCACTTTTCC 660
Qy 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Qy 721 TACTTAGTAATAAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGATGCTGTCAAA 780
Db 721 TACTTAGTAATAAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGATGCTGTCAAA 780
Qy 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCCAGTTGGCTCATTTGGAATGAATGCAAT 840
Db 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCCAGTTGGCTCATTTGGAATGAATGCAAT 840
Qy 841 TTGATGAACACAGTACATTTAGTTGTAGCTGACAGATTACTTTGTGGAACCTTCGATTTCA 900
Db 841 TTGATGAACACAGTACATTTAGTTGTAGCTGACAGATTACTTTGTGGAACCTTCGATTTCA 900
Qy 901 AAGGTTTTTCAGGCAGAAAATCCTTTGATTTTATGGAACAATTTCTTTAGAAGGAAA 960
Db 901 AAGGTTTTTCAGGCAGAAAATCCTTTGATTTTATGGAACAATTTCTTTAGAAGGAAA 960
Qy 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTGGAGTTATGCGAGAAACC 1020
Db 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTGGAGTTATGCGAGAAACC 1020
Qy 1021 ACAGATAAGCTCTTCCACCTTGGATCGAGATTTT 1053
Db 1021 ACAGATAAGCTCTTCCACCTTGGATCGAGATTTT 1053
```

RESULT 6

```
US-10-698-228-4
; Sequence 4, Application US/10698228
; Publication No. US2004007253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619W00P
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
```

```
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 4
; LENGTH: 1081
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-4
```

Query Match 99.8%; Score 1051.4; DB 19; Length 1081;

Best Local Similarity 99.9%; Pred. No. 3.le-294; Indels 0; Gaps 0;

Matches 1052; Conservative 0; Mismatches 1;

```
Qy 1 ATGGCGACCCCGGAAAGCGGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 20 ATGGCGACCCCGGAAAGCGGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 79
Qy 61 GACACCAACGAAAGTGAATTAAGTCAAATGAAGAGCCACTCCTAAGAAAGAGTTCTGC 120
Db 80 GACACCAACGAAAGTGAATTAAGTCAAATGAAGAGCCACTCCTAAGAAAGAGTTCTGC 139
Qy 121 CGSTTTGTCTATCTTCCAAATCCAGTACCCTGATATTTGGAAATATGATAACAGSCACAG 180
Db 140 CGSTTTGTCTATCTTCCAAATCCAGTACCCTGATATTTGGAAATATGATAACAGSCACAG 199
Qy 181 GCTTCCTCTTGGACAGCAGAGAGGTGCGACTTATCAAGAGATCTCCCTCACTGGAACAAG 240
Db 200 GCTTCCTCTTGGACAGCAGAGAGGTGCGACTTATCAAGAGATCTCCCTCACTGGAACAAG 259
Qy 241 CTTAAAGCAGATGAGAGTACTTCACTCTCAATCTTAGCCTTTTTCAGCCAGTGAT 300
Db 260 CTTAAAGCAGATGAGAGTACTTCACTCTCAATCTTAGCCTTTTTCAGCCAGTGAT 319
Qy 301 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGCT 360
Db 320 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGCT 379
Qy 361 CCGTGTCTATGCGCTTTCAAAATCTCATCGAATGTTTCACTCAGAGATGATCAGTTTG 420
Db 380 CCGTGTCTATGCGCTTTCAAAATCTCATCGAATGTTTCACTCAGAGATGATCAGTTTG 439
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 480
Db 440 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 499
Qy 481 ACCATGCCCTATGTTAAAGAAAGAGAGATTGGGCTTCGATGGATAGAGATAGAAA 540
Db 500 ACCATGCCCTATGTTAAAGAAAGAGAGATTGGGCTTCGATGGATAGAGATAGAAA 559
Qy 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTGTAGAGAGGATTTTCTTCAGGA 600
Db 560 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTGTAGAGAGGATTTTCTTCAGGA 619
Qy 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGCTTTATGCGAGGACTCACATTTTCC 660
Db 620 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGCTTTATGCGAGGACTCACATTTTCC 679
Qy 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA 720
Db 680 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA 739
Qy 721 TACTTAGTAATAAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGATGCTGTCAAA 780
Db 740 TACTTAGTAATAAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGATGCTGTCAAA 799
Qy 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCCAGTTGGCTCATTTGGAATGAATGCAAT 840
Db 800 ATTGAGCAGGAGTTTAAACAGAGCCTTGCCAGTTGGCTCATTTGGAATGAATGCAAT 859
```


Db 1426 GAGTAAGAGAAATAATCAATGCTGTTCGATAGAAACAGGAGTTCCCTCACTGAGGCT 1367
Qy 809 TCCAGTTGGCTCATTTGAATGAATGATTTGATGAAACAGTACATTTGATTTGAG 868
Db 1366 TCCCTGTGAAGCTCATTTGGGATGAATGACCTTAATGAAGCAATACATTTGATTTGGG 1307
Qy 869 CTGACAGATTACTTTGGAACTTGGATTCTCAAGGTTTTCAGGCGAGAAAATCCCTTTTG 928
Db 1306 CAGACAGACTTATGCTGGAACTGGGTTTTCAGAGTTTTCAGATGAGAACCCATTG 1247
Qy 929 ATTTATGGAACAATTTCTTTAGAGGAAACAAAATTTCTTTGAGAAACGAGTTTCAG 988
Db 1246 ACTTTATGAGAATAATTTCACTGGAAGGAAAGACCTAACTCTTTGAGAGAGAGTAGGCG 1187
Qy 989 AGTATCAGCGTTTTCAGTATTGGCAGAAACACAGATAAGCTTTCACCTTGGATGAG 1048
Db 1186 AGTATCAGAGGATGGAGTGTCAAGTCCAAACAGAGAATTTCTTTTACCTTTGGATGCTG 1127
Qy 1049 ATTT 1052
Db 1126 ACTT 1123

RESULT 10
US-09-954-456-724
; Sequence 724, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
; FILE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 724
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-724

Query Match 54.9%; Score 578.4; DB 9; Length 2500;
Best Local Similarity 75.0%; Pred. No. 1.4e-156;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;
Qy 89 ATGAAGAGCCACTCCTAAGAAAGATTCTCGCGGTTTGTCTATCTTTCCATCCAGTACC 148
Db 397 AGGATGAGCCGCTGCTGAGAGAAAACCCCGCGCTTTGTCTATCTTTCCCATCGAGTACC 456
Qy 149 CTGATATTTGGAAATGTATAACAGGCACAGGCTTCTCTTGGACAGAGAGGTCG 208

Db 457 ATGATATCTGGCAGATGTATTAAGAGGCAGAGGCTTCTTTTGGACCGCGAGAGGTTG 516
Qy 209 ACTTATCAAGAGATCTCCCTCACTGGAACAAGCTTAAGCAGATGAGAAGTACTTCACT 268
Db 517 ACCTCTCCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT 576
Qy 269 CTCACATCTTAGCCTTTTGGACCGCAGTGTGAATTTGAAATGAAAAATTTGGTGAGC 328
Db 577 CCCATGTTCTGGCTTTCTTTGCAAGCAAGCGATGCGATAGTAATGAAACTTGGTGAGC 636
Qy 329 GCTTTAGTCAGAGGTGCGAGTTCCAGAGGCTCGCTGTTCTATGCTTTTCAATTTCTCA 388
Db 637 GATTTAGCCAAGAGTTTCAGATTACAGAAGCCGCTGTTTCTATGCTTCCAAATTTGCCA 696
Qy 389 TCGAGAACTCTCACTCAGAGATGTACAGTTTGTCTAGTACACTTACATCAGAGATCCCA 448
Db 697 TGAAGAACATACATCTGAAATGTATAGTCTTCTTATGACACTTACATAAAGATCCCA 756
Qy 449 AGAAAGGGAATTTTATTTAATGCAATTTGAAACCATGCGCTATGTTTAAAGAAAAAGCAG 508
Db 757 AAGAAGGGAATTTCTTTCAATGCCATTTGAAGAGATGCCCTTGTCTCAAGAGAGCAG 816
Qy 509 ATTGGGCTTGGCAGTATGAGATGAGATAAGAAATCTACTTTTGGGAAAGAGTGGTGCT 568
Db 817 ACTGGGCTTGGCTGGATTTGGGACAAAGAGGCTACCTATGCTGAACGTTGTGTAGCCT 876
Qy 569 TTGCTGCTAGAGAGTTTCTTCTCAGGATCTTTGCTGCTATATCTGCTAAGTAAAGA 628
Db 877 TTGCTGCTAGTGAAGGCAATTTCTTTTCGGGTTCTTTTGGCTGATATCTGCGCTCAAGA 936
Qy 629 AGAGAGGCTTATGCCAGGACTCACATTTTTCATGAACTCATCAGCAGAGATGAAGAC 688
Db 937 AAGAGGACTGATGCTGGCTCACATTTTCTAATGAATTTATGAGAGATGAGGTT 996
Qy 689 TTCACGTGCACTTTGCTTGGCTGATGTTCAATCTTAGTAAATAAGCCTTTCAGAGAAA 748
Db 997 TACACTGTGATTTGCTTGGCTGATGTTCAAAACACCTGGGTACACAAACCATCGAGGAGA 1056
Qy 749 GGTGAGGAGATCATTTGATGCTGCTCAAAATGAGCAGGAGTTTTCACAGAACCT 808
Db 1057 GAGTAAGAGAAATAATTAATCAATGCTGTTGCGATAGAACAGGAGTTTCTCCTGAGGCT 1116
Qy 809 TGCCAGTTGGCCTCATTTGGAATGAAATGCAATTTTGTATGAAACAGTACATTTGATTTGAG 868
Db 1117 TGCCTGTGAGCTCATTTGGATGAATTTGCACTTAATGAGCAATACATTTGATTTGTTG 1176
Qy 869 CTGACAGATTACTTTGGAACTTTGGATTCTCAAGGTTTTCAGGCGAGAAAATCCTTTG 928
Db 1177 CAGACAGACTTATGCTGGAACCTGGGTTTTCAGCAAGGTTTTCAGAGTAGAGAACCAATTG 1236
Qy 929 ATTTATGGAACAATTTCTTTAGAGGAAACAAAATTTCTTTTGAACAGGATTTTCAG 988
Db 1237 ACTTTATGGAATAATTTCTGAGAGGAAAGACTAACTTTCTTTGAGAGAGATGAGCG 1296
Qy 989 AGTATCAGCGTTTTCAGTATTGCGCAGAAACACAGATAACGTTCTTCACTTGGATGAG 1048
Db 1297 AGTATCAGGATGGAGTGAUTCAAGTCCAAACAGAGAATTTCTTTTACCTTGGATGCTG 1356
Qy 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 11
US-09-954-456-1169
; Sequence 1169, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
; FILE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456

```

CURRENT FILING DATE: 2001-09-18
PRIOR APPLICATION NUMBER: US/60/233,617
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: US/60/234,052
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,923
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,134
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,637
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,638
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,711
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,720
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,840
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,863
PRIOR FILING DATE: 2000-09-27
NUMBER OF SEQ ID NOS: 2276
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1169
LENGTH: 2500
TYPE: DNA
ORGANISM: Homo sapiens
US-09-954-456-1169

```

Query Match	54.9%	Score 578.4	DB 9	Length 2500	
Best Local Similarity	75.0%	Pred. No. 1.4e-156			
Matches 723	Conservative	0	Mismatches 241	Indels 0	Gaps 0
QY	89	ATGAAGAGCCACTCCTAAAGAAAGAGTTC	CGCGGTTTGT	CATCTTTCCAAATCCAGTACC	148
DB	397	AGGATGAGCGCTGCTGAGNAAAA	CCCCCGCGCTTTGTCA	TCTCCCATCGATACC	456
QY	149	CTGATATTTGGAAAAATGTTATAA	CAGGCACAGGCTTCTTT	CTGSACACGAGAAGAGTCG	208
DB	457	ATGATATCTGGCAGATGTATAA	AGGCGCAGAGGCTTCTTT	TGGACCGCGAGGAGTGTG	516
QY	209	ACTTATCAAAAGGATCCTCCTCA	CTCGNAAACAAAGCTTAA	AGCAGATGAGAAAGTACTTCA	268
DB	517	ACCTCTCCAAGGACATTCAGCA	CTGGAAATCCCTGAAAC	CCGAGGAGAGATATTTTATAT	576
QY	269	CTCACATCTTAGGCTTTTTCG	ACGCCAGTCGATGGAAT	TGTAATGAAAAATTTGGTGGAGC	328
DB	577	CCCATGTTCTGGCTTTCTTT	TGCAGCAAGCGATG	GCATAGTAATGAAAACTTGGTGGAGC	636
QY	329	GCTTTAGTCAGGAGGTGCAGG	TTCCAGAGGCTCGCTG	TTTCTATGGCTTTCAAATTCATCA	388
DB	637	GAATTTAGCCAAGAAGTTCA	GATTTACAGAACCCGCT	GTTTTCTATGGCTTCCAAATTTGCCA	696
QY	389	TCGAGAAATGTCTACTCAG	AGATGTACAGTTTGTGT	GATAGACACTTACATCAGAGATCCCA	448
DB	697	TGGAAAAACATACATCT	GAAATGATATAGTCTT	TATGTACACTTACATTAAGATCCCA	756
QY	449	AGAAAAAGGAAATTTTAT	TTTAATGCAATTAAGAA	ACCATGCCCTATGTTTAAGAAAAAAGCAG	508
DB	757	AAGAAAGGGAATTTCTCT	TTCAATGCCATTTGAA	ACGATGCCCTGTGTCAAGAAAGAGGAGC	816
QY	509	ATTGGGCTTGGATGGATAG	CAGATAGAAAAATCTA	CTTTTGGGAAAGAGTGGTGGCTT	568
DB	817	ACTGGGCTTTGGCTGGAT	TGGGGAACAAGAGGCT	ACCTATGGTGAACGTTGTGTAGCCTT	876
QY	569	TTGCTGCTAGAAAGAGT	TTTCTCTCAGGATCT	TTTTTGTGCTATATTTCTGGCTAAAGA	628
DB	877	TTGCTGCTAGTGAAGGCA	TTTTTCTTTTCCGGT	TCTTTTGGCTCGATATTTCTGGCTCAAGA	936
QY	629	AGAGAGGTCTTATGCCAG	CACTCACTTTTTC	CAATGAACTCATCAGCAGAGATGAAGGAC	688
DB	937	AACAGGACGTAGTCCTG	CGCCTCACATTTTCT	TAATGAACCTATTATAGCAGAGATGAGGGTT	996

RESULT 12

```

US-09-954-456-1827
; Sequence 1827, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1827
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-1827

```

Query Match 54.9%; Score 578.4; DB 9; Length 2500;

Best Local Similarity 75.0%; Pred. No. 1.4e-156;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Caps 0;

89 ATGAAGAGCCACTCCTAAGAAAGAGTTCTCGCCGGTTTGTCATCTTTCCAAATCCAGTACC 148

Qy	689	TTCACTGTGACTTTGCTTGCCTGATGCTTCAATCACTTGTAGTAAATPAGCCCTTCAGAAGAAA	748
Db	997	TACACTGTGATTTTGTCTTGCCTGATGTTCAACACACCTGGTACACAAACCATCGGAGGAGA	1056
Qy	749	GGGTCAGGGAGATCATTTGTGATGCTGTCAAAATTGAGCAGGAGTGTTTTAAACAGACGCT	808
Db	1057	GAGTAAAGAGAAATAATTATCAATGCTGTTCGGATAGAACAGGAGTTCCTCACTGAGGCCT	1116
Qy	809	TGCCAGTTGCGCCTCATTTGGAAATGAATTGCAATTTTGATGAAACAGTACATTTGAGTTTGTAG	868
Db	1117	TGCTGTGAGCTCATTTGGATGNAATTGCACTCTAATGAAGCAATACATTTGAGTTTGTGG	1176
Qy	869	CTGACAGATTACTTTGTGGAACTTTGGATTCTCAAGGTTTTTCAGCGAGAAAATCCTTTTG	928
Db	1177	CAGACACACTTATGCTGGAACTGGGTTTTTAGCAAGGTTTTTCAGAGTAGAGAAACCCATTTG	1236
Qy	929	ATTTTATGGAAACATTTCTTTTAGAGGAAAAACAAATTTCTTTTGAAACGAGTTTTCAG	988
Db	1237	ACTTTATGGAGAATATTTTCACTCGAAGGAAGAAGCTAACTTCTTTTGAAGAGAGATGAGCG	1296
Qy	989	AGTATCAGCGCTTTTTCAGCTTTATGGCAGAAACCAAGATAACGCTTTCACCTTGGATGCGAG	1048
Db	1297	AGTATCAGAGGATGGGAGTGATGCAAGTCCACAGAGAATTCCTTTTACCTTGGATGCTG	1356
Qy	1049	ATTTT 1052	
Db	1357	ACTTT 1360	

RESULT 12
US-09-954-456-1827
; Sequence 1827, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1827
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-1827

Query Match 54.9%; Score 578.4; DB 9; Length 2500;

Best Local Similarity 75.0%; Pred. No. 1.4e-156;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Caps 0;

89 ATGAAGAGCCACTCCTAAGAAAGAGTTCTCGCCGGTTTGTCATCTTTCCAAATCCAGTACC 148

```
Db 397 AGGATGAGCGCTGCTGAGAGAAAACCCCGCGCTTTGTCTATCTTCCCATCGAGTACC 456
Qy 149 CTGATATTGGAAATGTATAACAGGCACAGGCTTCTTCTGGACAGAGAGGTCG 208
Db 457 ATGATATCTGGCAGATGTATAAGAGCGCAGAGGCTTCTTTTGGACCGCGAGGAGTTG 516
Qy 209 ACTTATCAAGAGATCTCCCTCACTGAGCAAGCTTAAAGCAGATGAGAAGTACTTCACT 268
Db 517 ACTCTCCAGAGACATTGAGCATCTGGATCCCTGAAACCGAGGAGATATTTATAT 576
Qy 269 CTCACATCTTAGCTTTTGGCAGCAGTGATGGAATTTAAATGAAAATTTGGTGAGC 328
Db 577 CCCATGTTCTGGCTTCTTTGACGCAAGCGATGGCATAGTAATGAAAATCTTGGTGAGC 636
Qy 329 GCTTTAGTCAGAGGTCGAGGTTCCAGAGGCTCGCTGTTCTATGGCTTTCCAAATCTCA 388
Db 637 GATTTAGCCAAAGATTTTCTTTGACGCAAGCGATGGCATAGTAATGAAAATCTTGGTGAGC 636
Qy 329 GCTTTAGTCAGAGGTCGAGGTTCCAGAGGCTCGCTGTTCTATGGCTTTCCAAATCTCA 388
Db 637 GATTTAGCCAAAGATTTTCTTTGACGCAAGCGATGGCATAGTAATGAAAATCTTGGTGAGC 636
Qy 389 TCAGAAATGTTTCACTCAGAGATGTACAGTTTGTGTAGTAGACACTTACATCAGAGATCCCA 448
Db 697 TGGAAAACATACATTTCTGAAATGTATAGTCTTCTTATTTGACACTTACATAAAGATCCCA 756
Qy 449 AGAAAAGGGAATTTTATTAATGCAATTTGAAACCAATGCCCTATGTTAAGAAAAAGAGCAG 508
Db 757 AAGAAAAGGGAATTTCTTCAATGCAATTTGAAACCAATGCCCTATGTTAAGAAAAAGAGCAG 816
Qy 509 ATTGGGCTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 568
Db 817 ACTGGGCTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 876
Qy 569 TTGCTGCTGATGAGAGGATTTTCTTCTCAGGATCTTTTCTGCTATATTTCTGCTAAAGA 628
Db 877 TTGCTGCTGATGAGAGGATTTTCTTCTCAGGATCTTTTCTGCTATATTTCTGCTCAAGA 936
Qy 629 AGAGAGGCTTTATGCGAGGACTCACCTTTTCCAATGAATCTATCAGCAGAGATGAAGGAC 688
Db 937 AACGAGGACTGATGCTGCGCTCACATTTTCTTAATGAATCTATCAGCAGAGATGAGGTT 996
Qy 689 TTCACCTGACATTTGCTTGCCTGATGTTCCAAATGCAATCTAGTAATAGCCTTCAGAGAAA 748
Db 997 TACACTGTGATTTGCTTGCCTGATGTTTCAAAACACCTGTTACAAACCATCGGAGGAGA 1056
Qy 749 GGGTCAGGAGATCATTTGTTGATGCTGTCAAAATTCAGCAGGAGTTTAAACAGAGGCT 808
Db 1057 GAGTAGAGAAATATATCAATGCTTGGTGGATGATGATGATGATGATGATGATGATGATGAT 1116
Qy 809 TGCCAGTTGGCTCATTTGGAATGAATTTGATGAAACAGTACATTTGAGTTTGTAG 868
Db 1117 TGCTGTGAAGCTCATTTGGGATGAATTTGACATCTATGAAGCAATACATTTGAGTTTGTG 1176
Qy 869 CTGACAGATTAATTTGTTGAACTTTGATTTCTCAAGGTTTTCAGGAGAAAATTCCTTTTG 928
Db 1177 CAGACAGACTTATGCTGGAATCTGGGTTTATGCAAGGTTTTCAGAGTAGAGAACCCATTG 1236
Qy 929 ATTTTATGAAAACATTTCTTTAGAGGAAAACAAATTTCTTGAAGAACAGGTTTCAG 988
Db 1237 ACTTTATGGAATATTTTCACTGGAAGGAAAGACTAACTTCTTTGAGAGAGAGTAGGCG 1296
Qy 989 AGTATCAGGCTTTTGCAGTTATGGCAGAAACACAGATAACGCTTCTTCACTTTGGATGCGAG 1048
Db 1297 AGTATCAGAGGATGGGATGATGTCAGTCCAAACAGAGATTTCTTTTACCTTGGATGCTG 1356
Qy 1049 ATTT 1052
Db 1357 ACTT 1360
```

RESULT 13

US-10-641-643-1370
; Sequence 1370, Application US/10641643
; Publication No. US20040077003A1
; GENERAL INFORMATION:

APPLICANT: Cocks, Benjamin G.
Susan G. Stuart
Jeffrey J. Seilhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
GENE EXPRESSION
NUMBER OF SEQUENCES: 1508
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/641,643
FILING DATE: 14-Aug-2003
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0001 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1370:
SEQUENCE CHARACTERISTICS:
LENGTH: 2500 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: G36154
SEQUENCE DESCRIPTION: SEQ ID NO: 1370 :
US-10-641-643-1370

Query Match 54.9%; Score 578.4; DB 19; Length 2500;
Best Local Similarity 75.0%; Pred. No. 1.4e-156;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

Qy	89	ATGAAGAGCCACTCCTTAAGAAAGAGTTCTCGCGGTTTGTCTATCTTTCCAAATCCAGTACC	148
Db	397	AGATGAGCGCTGCTGAGAGAAAACCCCGCGCTTTGTCTATCTTTCCCATCGAGTACC	456
Qy	149	CTGATATTGGAAAATGTATAACAGGCACAGGCTTCTTCTTGGACAGAGAGGTCG	208
Db	457	ATGATATCTGGCAGATGTATAAGAGCGCAGAGGCTTCTTTTGGACCGCGAGGAGTTG	516
Qy	209	ACTTATCAAGAGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAAGTACTTCACT	268
Db	517	ACCTCTCCAAGGACATTCAGACTCTGGGAATCCCTGAAAACCCGAGGAGATATTTTATAT	576
Qy	269	CTCACATCTTAGCTTTTGGCAGCCAGTGATGGAATTTGAAATGAAAATTTGGTGAGC	328
Db	577	CCATGTTCTGGCTTTCTTTGACGCAAGCGATGGCATAGTAATGAAAATCTTGGTGAGC	636
Qy	329	GCTTTAGTCAGAGGTCGAGGTTCCAGAGGCTCGCTGTTCTATGGCTTTCCAAATCTCA	388
Db	637	GATTTAGCCAAAGATTTTCTTTGACGCAAGCGCTGTTCTATGGCTTTCCAAATTTGCA	696
Qy	389	TCAGAAATGTTTCACTCAGAGATGTACAGTTTGTGTAGTAGACACTTACATCAGAGATCCCA	448
Db	697	TGGAAAACATACATTTCTGAAATGTATAGTCTTCTTATTTGACACTTACATAAAGATCCCA	756
Qy	449	AGAAAAGGGAATTTTATTAATGCAATTTGAAACCAATGCCCTATGTTAAGAAAAAGCAG	508

```
Db 757 AAGAAAGGGAATTTCTCTCAATGCGCAATGAAACGATGCGCTTGTGTCAAGAAAGAGCGAG 816
Qy 509 ATTGGGCGCTTGGGATAGCAGATAGAAATCTACTTTTGGGGAAGAGTGGTGGCCT 568
Db 817 ACTGGGCGCTTGGGATAGGGAACAAGAGGCTTACCTATGTTGAACGTTGTGTAGCCT 876
Qy 569 TTGCTGCTGTAGAAGGAGTTTCTCTCAGGATCTTTTGGCTGTATATTTCTGGCTAAAGA 628
Db 877 TTGCTGCACTGGAAGGCAATTTCTTTCCGGTCTTTTGGTGGATATTTCTGGCTCAAGA 936
Qy 629 AGAGAGGCTTTATGCCAGGACTCACTTTTCCAAATGAATCACTACAGCAGATGAAGGAC 688
Db 937 AACGAGGACTGATGCGCTGCGCTCACATTTTCTAAATGAATCTTATAGCAGATGAGGGTT 996
Qy 689 TTCACCTGTCACTTTGCTGCGCTGATGTTCCAAATGCTTAGTAATAGCCTTCAGAGAAA 748
Db 997 TACACTGTGATTTTGTCTGCGCTGATGTTCAAAACACCTGTTACAAACCATCGGAGGAGA 1056
Qy 749 GGGTCAGGAGAGATCAATTTGTGATGCTGTCAAAATTTGAGCAGGAGTTTAAACAGAGCCT 808
Db 1057 GAGTAAGAGAAATATATCAATGCTGTTGCGATAGAACAGAGTTCTCTCAGTGGCCT 1116
Qy 809 TGGCAGTTGGCTCACTTGAATGAATTTGATTTGATGAACAGTACATTTGAGTTGTAG 868
Db 1117 TGCCTGTGAAGCTCAATGGGATGAATTTGCACTCTAATGAAGCAATACATTTGAGTTGTGG 1176
Qy 869 CTGACAGATTACTTTGGNACTTGGATTTCTCAAGTTTTTCAGGAGAAAATCCTTTTG 928
Db 1177 CAGACAGACTTATGCTGGAACTGGGTTTGTAGCAAGGTTTTTCAGAGTAGAAGACCCATTTG 1236
Qy 929 ATTTTATGAAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAACGAGTTTTCAG 988
Db 1237 ACTTTATGGAGAAATTTCTGGAAGGAAAGACTAACTTTCTTTGAGAAGAGTAGGCG 1296
Qy 989 AGTATCAGCGTTTTCAGTTATGGCAGAAACACAGATAACGCTCTTCACTTTGGATGCGAG 1048
Db 1297 AGTATCAGAGGATGGGAGTGATGTCAAGTCCCAACAGAGAAATCTTTTACCTTGGATGCTG 1356
Qy 1049 ATTT 1052
Db 1357 ACTT 1360
```

RESULT 14

```
US-10-733-878-458
; Sequence 458, Application US/10733878
; Publication No. US20040224408A1
; GENERAL INFORMATION:
; APPLICANT: Jean-Philippe Girard
; APPLICANT: Francois Anallric
; APPLICANT: Myriam Roussigne
; APPLICANT: Thomas Clouaire
; TITLE OF INVENTION: THAP PROTEINS AS NUCLEAR RECEPTORS FOR
; TITLE OF INVENTION: CHEMOKINES AND ROLES IN TRANSCRIPTIONAL REGULATION, CELL
; TITLE OF INVENTION: PROLIFERATION AND CELL DIFFERENTIATION
; FILE REFERENCE: BIOBANK.012A
; CURRENT APPLICATION NUMBER: US/10/733,878
; CURRENT FILING DATE: 2003-12-10
; PRIOR FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 60/432699
; PRIOR FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/485027
; PRIOR FILING DATE: 2003-07-03
; NUMBER OF SEQ ID NOS: 535
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 458
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-733-878-458
```

Query Match

Best Local Similarity

54.9%; Score 578.4; DB 21; Length 2500;

75.0%; Pred. No. 1.4e-156;

RESULT 15

US-10-843-641A-3751

```
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;
Qy 89 ATCAAGAGCCACTCTCTAAGAAAGAGTTCTCGCGGTTTGTCTATCTTTTCAATTCAGTACC 148
Db 397 AGGATGAGCGCGCTGTGTGAGAGAAACCCCGCGCTTGTCTATCTTTCCCATTCAGTACC 456
Qy 149 CTGATATTTGGAATAATGTATATAACAGGCACAGGCTTCTCTTGTGACAGCAGAGAGTCTG 208
Db 457 ATGATATCTGCGAGATGTATAAGAAAGCAGAGGCTTCTTTTGGACCCCGGAGAGGTTG 516
Qy 209 ACTTATCAAAAGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAAGTACTTCTATCT 268
Db 517 ACCTCTCCAAGGACATTCAGCACTCTGGGAATCCCTGAAACCCGAGAGAGATATTTTATAT 576
Qy 269 CTCACATCTTAGCCTTTTGTGAGCCAGTATGGAATTTGTAATGAAATTTGTGGAGC 328
Db 577 CCCATGTTCTGGCTTCTTTTGCAGCAAGCGATGECATAGTAAATGAAACCTTGTGGAGC 636
Qy 329 GCTTTAGTCAGGAGGTGCGAGTTCCAGAGGCTCGCTTCTTCTATGGCTTTCAAAATTTCTCA 388
Db 637 GATTTAGCCAAAGATTCAGATTACAGAGCCCGCTGTTTCTATGGCTTCCAAATTTGCCA 696
Qy 389 TCAGAAATGTTCACTCAGAGATGTACAGTTTGTGTGATAGACACTTACATCAGAGATCCCA 448
Db 697 TGGAAACATACATCTTGAATGTAGTCTTCTTATTGACACTTACATAAAGATCCCA 756
Qy 449 AGAAAGGGAATTTTATTAATGCAATTTGAAACCATGCGCTATGTTAAGAAAAAGCAG 508
Db 757 AAGAAAGGGAATTTCTCTCAATGCGCAATGAAACGATGCGCTTGTGTCAAGAAAGGCGAG 816
Qy 509 ATTGGGCGCTTGGCATGATAGCAGATAGAAATCTACTTTTGGGGAAGAGTGTGGCCT 568
Db 817 ACTGGGCGCTTGGCTGATTTGGGGAACAAGAGGCTACCTATGTTGAACGTTGTAGCCT 876
Qy 569 TTGCTGCTGTAGAAGGAGTTTCTTCTCAGGATCTTTTGTGCTATATTTCTGGCTAAAGA 628
Db 877 TTGCTGAGTGAAGGCAATTTCTTTTCCGGTCTTTTGTGCTCGATATTTCTGCTCAAGA 936
Qy 629 AGAGAGTCTTATGCCAGGACTCACTTTTTCGAATGAACTCATCAGCAGAGATGAAGGAC 688
Db 937 AACGAGGACTGATGCGCTGCGCTCACATTTTCTAATGAACCTTATTAGCAGAGATGAGGGTT 996
Qy 689 TTCACCTGTGACTTTGCTGCGCTGATGTTCCAAATGCTTAGTAATAGCCTTCAGAAAGAA 748
Db 997 TACACTGTGATTTTGTCTGCGCTGATGTTCAACACCTGGTACAAACCATCGGAGGAGA 1056
Qy 749 GGGTCAGGAGAGATCAATTTGTGATGCTGTCAAAATTTGAGCAGGAGTTTAAACAGAGCCT 808
Db 1057 GAGTAAGAGAAATATATCAATGCTGTTGCGATAGAACAGGAGTTCTCTCAGTGGCCT 1116
Qy 809 TGGCAGTTGGCTCACTTGAATGAATTTGATTTGATGAACAGTACATTTGAGTTGTAG 868
Db 1117 TGCCTGTGAAGCTCAATGGGATGAATTTGCACTCTAATGAAGCAATACATTTGAGTTGTGG 1176
Qy 869 CTGACAGATTACTTTGGGAGCTTGGATTTCTCAAGGTTTTTTCAGGAGAAAATCCTTTTG 928
Db 1177 CAGACAGACTTATGCTGGAACTGGGTTTGTAGCAAGGTTTTTCAGAGTAGAAGACCCATTTG 1236
Qy 929 ATTTTATGAAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAACGAGTTTTCAG 988
Db 1237 ACTTTATGGAGAAATTTCTGGAAGGAAAGACTAACTTTCTTTGAGAAGAGTAGGCG 1296
Qy 989 AGTATCAGCGTTTTCAGTTATGGCAGAAACCAAGATAACGCTCTTCACTTTGGATGCGAG 1048
Db 1297 AGTATCAGAGGATGGGAGTGATGTCAAGTCCCAACAGAGAAATCTTTTACCTTGGATGCTG 1356
Qy 1049 ATTT 1052
Db 1357 ACTT 1360
```

; Sequence 3751, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals, Inc.
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; FILE REFERENCE: 689290-189
; CURRENT APPLICATION NUMBER: US/10/843,641A
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/873,367
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US/09/954,531
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,832
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/964,824
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/09/967,768
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/09/968,007
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,708
; PRIOR FILING DATE: 2001-10-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 8447
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3751
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-843-641A-3751

Query Match 54.9%; Score 578.4; DB 22; Length 2500;
Best Local Similarity 75.0%; Pred. No. 1.4e-156;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;
QY 89 ATGAAGAGCACTCCCTAAGAAAGAGTTCTCGCGGTTTGTCTCATCTTCCAAATCCAGTACC 148
DB 397 AGGATGAGCGCTGTGAGAGAAACCCCGCGGCTTGTCTCATCTTCCCATCGAGTACC 456
QY 149 CTGATATTTGGAAATGTATAACAGGCACAGGCTTCTTCTTGACAGCAGAGAGGTG 208
DB 457 ATGATATCTGGCAGATGTATAAGAGGCAGAGGCTTCTTCTTGGACCGCGAGGAGTTG 516
QY 209 ACTTATCAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAAGTACTTCATCT 268
DB 517 ACCTCTCAAGGACATTTCAAGCATGGGAAATCCCTGAAACCCGAGGAGAGATATTTATAT 576
QY 269 CTCACATCTTAGCTTTTGGACGAGTGAGGAAATGTAATGAAATTTGGTGGAGC 328
DB 577 CCCATGTTCTGGCTTCTTTGACGAGCAGATGGCATAGTAATGAAACTTGGTGGAGC 636
QY 329 GCTTTAGTCAGGAGTGCAGGTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAAATCTCA 388
DB 637 GATTTAGCCAAGAGTTTCAGATTACAGAAGCCGCTGTTTCTATGGCTTCCAAATTTGCCA 696
QY 389 TCGAGAATGTTCACTCAGAGATGTACAGTTTGTGTAGTACACACTTACATCAGAGATCCCA 448
DB 697 TGGAAAAACATACATCTCGAAATGTATAGTCTTCTTATTGACACTTACATAAAGATCCCA 756
QY 449 AGAAAGGGAATTTTATTAATGCAATGAAACCATGCCCTATGTTTAAAGAAAAAGCAG 508
DB 757 AAGAAAGGGAATTTCTCTTCAATGCAATGAAACCATGCCCTTGTGTCAAGAAAGAGGCAG 816
QY 509 ATTGGGCTTGGGATGGATAGAGATAGAAAAATCTACTTTTGGGAAAGAGTGGTGGCT 568
DB 817 ACTGGGCTTGGCTGGATTTGGGCAAGAGAGGCTACCTATGCTGAACGTTGTTAGCCT 876

QY 569 TTGCTGCTGTAGAAAGAGTTTTCTTCTCAGGATCTTTTGTCTATATTTCTGGCTAAAGA 628
DB 877 TTGCTGCTGAGGAGCAATTTCTTTCCGTTCTTTTGTCTGATATTTCTGGCTCAGA 936
QY 629 AGAGAGTCTTATGCCAGGACTCACATTTTCCAATGAATCATCAGCAGAGATGAAGAC 688
DB 937 AACGAGGACTGATGCTGCGCTCACATTTTCTAATGAACCTTATTAGCAGAGATCAGGTT 996
QY 689 TTCACCTGTGACTTTGCTTGGCTGATGTTCCAAATCTTAGTAATTAAGCCTTCAGAA 748
DB 997 TACACTGTGATTTGCTTGGCTGATGTTCAAAACACCTGGGTACACAAACCATCGAGG 1056
QY 749 GGCTCAGGAGGATCATTTGTTGATGCTGTCAAAATTTGAGCAGGAGTTTTTAAACA 808
DB 1057 GAGTAAGAGAAATAATATCAATGCTGTTCCGATAGAACAGGAGTCTCTCACTGAG 1116
QY 809 TGCAGTTGGCCTCATTTGGAATGAATTTGATGAAACAGTACATTTGAGTTTGTAG 868
DB 1117 TGCCTGTGAAGCTCATTTGGGATGAATTTGCACTCTAATGAAGCAATACATTTGAG 1176
QY 869 CTGACAGATTTACTTTGGAACCTTGGATTTCTCAAGGTTTTTTCAGGCAAGAAATCCT 928
DB 1177 CAGACAGACTTATGCTGGAACCTGGGTTTTAGCAAGGTTTTTTCAGAGTAGAGAAC 1236
QY 929 ATTTTATGGAACATTTCTTTAGAAGGAAAAACAAATTTCTTTGAGAAACGAGTTTC 988
DB 1237 ACTTTATGAGAAATATTTCACTGGAAGGAAAGAACTAACTTCTTTGAGAAAGAT 1296
QY 989 AGTATCAGGCTTTTTCAGATTTATGCGCAGAAACCAAGATACGCTTTCACCTTGGAT 1048
DB 1297 AGTATCAGAGGATGGGAGTGTGTCAGTCCAAACAGAGAAATTTCTTTTACCTT 1356
QY 1049 ATTT 1052
DB 1357 ACTT 1360

Search completed: October 30, 2005, 06:24:09
Job time : 679.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 30, 2005, 03:23:39 ; Search time 62.25 Seconds
(without alignments)
13084.616 Million cell updates/sec

Title: US-10-698-228-2

Perfect score: 1846

Sequence: 1 atggggcaccggaaaggcc.....tcaccttgatgcagatttt 1053

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 4211384

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO spo01/US10698228/runat_26102005_100607_3977/app_query.fasta_1.2446
-DB=A_Geneseq_16Dec04 -QFMT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -ENDS=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10698228 @CGN_1_1_140 @runat_26102005_100607_3977 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_16Dec04.*

- 1: geneseqp1980s.*
- 2: geneseqp1990s.*
- 3: geneseqp2000s.*
- 4: geneseqp2001s.*
- 5: geneseqp2002s.*
- 6: geneseqp2003as.*
- 7: geneseqp2003bs.*
- 8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1821	98.6	351	3	Aay84439 Amino aci
2	1821	98.6	351	4	Aab69050 Human rib
3	1817	98.4	351	4	Aab93497 Human pro
4	1454	78.8	389	4	Aab69051 Human R2
5	1454	78.8	389	6	Abu07433 Protein d
6	1454	78.8	389	7	Adj68753 Human hea
7	1454	78.8	389	8	Adj66564 Ribonucle
8	1454	78.8	389	8	Adk70435 Respirato
9	1454	78.8	389	8	Adn03789 Antipsori
10	1454	78.8	389	8	Adn04444 Antipsori

11	1454	78.8	389	8	ADO19226	Ado19226 Human PRO
12	1454	78.8	389	8	ADQ09272	Adm09272 Human RRM
13	1454	78.8	389	8	Abm80258	Abm80258 Tumour-as
14	1454	78.8	413	3	AA433902	Human can
15	1454	78.8	453	4	AAU28017	Human con
16	1454	78.8	453	4	ABG15287	Novel hum
17	1429	77.4	390	2	AAW41765	Human rib
18	1425.5	77.2	386	8	ADO57309	Kidney de
19	1348	73.0	393	4	ABB64130	Abb64130 Drosophi
20	1242	67.3	430	8	ADN21469	Adn21469 Bacterial
21	1230	66.6	405	6	ABJ26187	ABJ26187 Aspergill
22	1214	65.8	381	8	ADN23033	Adn23033 Bacterial
23	1192.5	64.6	381	6	ABJ25587	ABJ25587 Aspergill
24	1192	64.6	413	5	ABP73700	Abp73700 Candida a
25	1176.5	63.7	321	8	ADS21332	AdS21332 Bacterial
26	1169.5	63.4	399	6	ABR53196	ABR53196 Protein s
27	1169.5	63.4	399	7	ADK63240	AdK63240 Disease t
28	1169.5	63.4	399	8	ADS43740	AdS43740 Bacterial
29	1161.5	62.9	391	8	ADS44135	AdS44135 Bacterial
30	1161.5	62.9	401	7	ADB70101	AdB70101 C. neofo
31	1132	61.3	352	3	AG05697	Ag05697 Arabidops
32	1130	61.2	329	3	AGS1528	AGS1528 Arabidops
33	1130	61.2	329	3	AG05699	Ag05699 Arabidops
34	1130	61.2	332	3	AGS1527	AGS1527 Arabidops
35	1130	61.2	332	3	AG05698	Ag05698 Arabidops
36	1117	60.5	347	3	AGI4546	AGI4546 Arabidops
37	1112.5	60.3	330	3	AGI4548	AGI4548 Arabidops
38	1112.5	60.3	333	3	AGI4547	AGI4547 Arabidops
39	1099	59.5	314	3	AGS1529	AGS1529 Arabidops
40	1091	59.1	341	3	AGG41395	AGG41395 Arabidops
41	1091	59.1	363	3	AGG41394	AGG41394 Arabidops
42	1084.5	58.7	330	3	AGG41396	AGG41396 Arabidops
43	1072.5	58.1	274	2	AAW26418	AAW26418 Swinepox
44	1072.5	58.1	274	4	AAW68239	AAW68239 Protein e
45	977.5	53.0	322	2	AAR72736	Aar72736 Plasmodiu

ALIGNMENTS

RESULT 1

AAV84439

ID AAV84439 standard; protein; 351 AA.

AC AAV84439;

DT 25-JUL-2000 (first entry)

DE Amino acid sequence of a human RNA-associated protein.

XX Human; RNA-associated protein; cell proliferation; cancer; inflammation;
XX Immune response; reproductive disorder; actinic keratosis;
XX atherosclerosis; arteriosclerosis; bursitis; cirrhosis; hepatitis;
XX mixed connective tissue disease; myelofibrosis; primary thrombocythemia;
XX paroxysmal nocturnal hemoglobinuria; polycythemia vera; psoriasis;
XX trauma.
XX Homo sapiens.
XX Key Location/Qualifiers
XX Modified-site 18 /note= "potential phosphorylation site"
XX Modified-site 23 /note= "potential glycosylation site"
XX Modified-site 29 /note= "potential phosphorylation site"
XX Modified-site 38 /note= "potential phosphorylation site"
XX Modified-site 39 /note= "potential phosphorylation site"
XX Modified-site 65 /note= "potential phosphorylation site"
XX Modified-site 87 /note= "potential phosphorylation site"

FT Modified-site 182 /note= "potential phosphorylation site"
 FT Modified-site 225 /note= "potential phosphorylation site"
 FT Modified-site 314 /note= "potential glycosylation site"
 FT Modified-site 329 /note= "potential phosphorylation site"

XX WO200015799-A2.

XX PD 23-MAR-2000.

XX PF 17-SEP-1999; 99WO-US021688.

XX PR 17-SEP-1998; 98US-00156039.

XX PR 22-SEP-1998; 98US-00158720.

XX PR 04-NOV-1998; 98US-00186815.

XX PR 08-APR-1999; 99US-0128660P.

XX PA (INCY-) INCYTE PHARM INC.

XX PI Tang YT, Corley NC, Guegler KJ, Gorgone GA, Patterson C;
 PI Hillman JL, Baughn MR, Lai P, Azimzai Y, Yue H, Yang J;
 XX N-PSDB; AAA12411.

DR WPI; 2000-271437/23.

DR N-PSDB; AAA12411.

XX New polypeptides and polynucleotides, useful for preventing and treating
 a disorder associated with increased or decreased expression of RNA
 associated proteins.

XX Claim 1; Page 103-104; 131pp; English.

XX The present sequence represents a human RNA-associated protein. The
 expression of RNA-associated proteins is closely associated with
 reproductive tissues, nervous tissues, cell proliferation including
 cancer, inflammation and immune responses, and so they may be used for
 diagnosis, treatment or prevention of cell proliferative,
 immune/inflammatory disorders, and reproductive disorders. Diseases and
 disorders which may be treated include actinic keratosis,
 atherosclerosis, arteriosclerosis, bursitis, cirrhosis, hepatitis, mixed
 connective tissue disease, myelofibrosis, paroxysmal nocturnal
 hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia
 and cancers, and trauma

XX Sequence 351 AA;

Alignment Scores:

Pred. No.: 1,03e-209 Length: 351

Score: 1821.00 Matches: 351

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 98.65% Indels: 0

DB: 3 Gaps: 0

US-10-698-228-2 (1-1053) x AAY84439 (1-351)

QY 1 ATGGCGACCCGGAAGCGCGCGGCTGGATCAGATCAGATCATCTTCA 60

DB 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSer 20

QY 61 GACACCAACGAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAAGTCTCGC 120

DB 21 AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerArg 40

QY 121 CGGTTTGTTCATCTTCCATCCAGTACCTGATATTGGAAATCTATAAAGAGGACAG 180

DB 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60

QY 181 GCTTCCTCTCGACAGCAGAGAGTGCAGTATCAAGAGATCTCCCTCACCTGGACACAG 240

DB 61 AlaSerPheTrpThrAlaGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80

QY 241 CTTAAAGCAGATCAGAGAGTACTTCACTCTCACATCTTACGCTTTTTCGAGCCAGTGAT 300
 DB 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
 QY 301 GGAATGTAAATCAAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTCCAGAGCT 360
 DB 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATGAGAAATGTTTCACTCAGAGATGTACAGTTG 420
 DB 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 QY 421 CTGATAGACACTTACATCAGATCCCAAGAAAAGGAAATTTTATTAATGCAATTCAG 480
 DB 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGCCCTTCGATGATGATGATGATGATG 540
 DB 161 ThrMetProTyrValLysLysLysAlaAspTyrAlaLeuArgTrpIleAlaAspArgLys 180
 QY 541 TCTACTTTTGGGAAAAGAGTGGTGGCTTTGCTGCTGTAGAGGAGTTCCTTCTCAGGA 600
 DB 181 SerThrPheGlyGluArgValValAlaPheAlaValGluGluValPhePheSerGly 200
 QY 601 TCTTTTGTCTGCTATATTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACATTTTCC 660
 DB 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 QY 661 AATGAACCTCATCAGCAGATGAGGACTTCCTCTGACCTTCTGCTGCTGCTGCTGCTG 720
 DB 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 QY 721 TACTTAGTAAATAGCCTTCAGAAAGAGAGTTCAGGAGATCATTTGTTGCTGTCTCAA 780
 DB 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
 QY 781 ATTGACGACGAGTCTTTTAAACAGAGAGCTTCGCCAGTGGCTCATTTGGAATGAATTC 840
 DB 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 QY 841 TTGATGAACACAGTACATTCAGTTGTTGATGCTGACAGATTAATCTGTGGAATCTGATCTCA 900
 DB 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 QY 901 AAGTTTTTTCAGCAGAAAATCCTTTTGTATTTTATGGAACAAACATTTCTTTAGAGGAAA 960
 DB 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 QY 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGTCAGTATG 1020
 DB 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 QY 1021 ACAGATAACGCTCTTCACTTGGATGACAGATTTT 1053
 DB 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 2
 ID AAB69050
 XX AAB69050 standard; protein; 351 AA.
 AC AAB69050;
 XX
 DT 18-APR-2001 (first entry)
 DE Human ribonucleotide reductase TP53R2H protein sequence SEQ ID NO:1.
 XX Human ribonucleotide reductase; cancer; DNA repair; p53.
 OS Homo sapiens.
 XX
 PN WO200100799-A1.
 XX

PD 04-JAN-2001.
 XX
 PF 27-JUN-2000; 2000WO-JP004189.
 XX
 PR 28-JUN-1999; 99JP-00181131.
 PR 06-JUL-1999; 99JP-00192391.
 PR 21-JAN-2000; 2000JP-0001770.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 PA (NAKA/) NAKAMURA Y.
 XX
 PI Nakamura Y, Arakawa H, Tanaka H;
 XX N-PSDB; AAF32438.
 DR WPI; 2001-112446/12.
 DR
 XX
 PT Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
 PT diagnosis, treatment and prevention of cancer.
 XX
 PS Claim 1; Fig 1-3; 102pp; Japanese.
 CC The present sequence represents a human ribonucleotide reductase
 CC designated TP53R2H. TP53R2H has cytosolic activity. TP53R2H is part of
 CC the DNA repair mechanism and its activity is induced by p53. It can be
 CC used for the treatment, prevention and diagnosis of a wide range of
 CC cancers
 CC
 XX SQ Sequence 351 AA;

Alignment Scores:
 Pred. No.: 1,03e-209 Length: 351
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 98.65% Indels: 0
 DB: 4 Gaps: 0

US-10-698-228-2 (1-1053) x AAB69050 (1-351)

QY 1 ATGGGGGAGCCCGAAGAGCGCGGCTGGATCAGATGAGAGATCATCTTCA 60
 DB 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSer 20
 QY 61 GACACCAAGCAAGTGAATTAAGTCAATGAGAGCCACTCTAAGAAGAGTTCTCC 120
 DB 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
 QY 121 CGGTTTGTCATCTTCCAAATCCAGTACCTGTATTTGAAAATGTATAAAGCAGCACAG 180
 DB 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
 QY 181 GTTCTCTTCTGACAGCAGAGAGTTCGACTTATCAAGAGTCTCCCTCACTGGACAAAG 240
 DB 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
 QY 241 CTTAAGCAGATGACAGTACTTCTCTCATCTTACATCTTACGCTTTTTCAGCCAGTAT 300
 DB 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
 QY 301 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGTTCAGAGGCT 360
 DB 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 QY 361 CGCTGTTTCTATGGCTTTCAAAATTCATCGAGAATGTTTCACTCAGAGATGTACAGTTTG 420
 DB 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
 DB 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 QY 481 ACCATGCCCTATGTTAAGAAAAACAGATTCGGCTTCGGATGCGATACAGATAGAAA 540

Db 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
 QY 541 TCTACTTTTGGGAAAGAGTGGCTTCTGCTGTAGAGGAGTCTTCTTCTCAGCA 600
 Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
 QY 601 TCTTTTCTGCTATATTCTGGCTAAAGAAGAGAGTCTTATGCCAGGACTTCACTTTTCC 660
 Db 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 QY 661 AATGAATCATCAGCAGAGATGAAGACTTTCACCTGTGACTTTGCTTGTGCTGTATGTTCAA 720
 Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 QY 721 TACTTAGTAATAAGCTTCCAGAGAAAGGTCAGGAGATCATTTGTTGATGCTCTCAA 780
 Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
 QY 781 ATTGAGCAGGAGTTTTAAACAGAGCCTTGCAGTTGGCCTCATTTGGAATGAATTCATT 840
 Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 QY 841 TTGATGAACAGTACATGAGTTTGTAGTCAACAGATTTACTTTGGAACCTTGGATTCTCA 900
 Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluGlyPheSer 300
 QY 901 AAGTTTTTCAGGACAGAAATCCTTTTGTATTTTATGAAACATTTCTTTAGAACGAAA 960
 Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 QY 961 ACAAAATTTCTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTATATGSCAGAAACC 1020
 Db 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 QY 1021 ACAGATAACGCTTTCACCTTGGATGCAATTTT 1053
 Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 3
 AAB93497
 ID AAB93497 standard; protein; 351 AA.
 XX
 AC AAB93497;
 XX
 DT 26-JUN-2001 (first entry)
 DE Human protein sequence SEQ ID NO:12811.
 XX
 DE Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX
 KW Homo sapiens.
 OS
 XX
 PN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-00116126.
 XX
 PR 29-JUL-1999; 99JP-00248036.
 PR 27-AUG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI; 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or

PS Example 2; Page 92-94; 102pp; Japanese.

XX The present invention describes a human ribonucleotide reductase
 CC designated TP53R2H. TP53R2H has cytosolic activity. TP53R2H is part of
 CC the DNA repair mechanism and its activity is induced by p53. It can be
 CC used for the treatment, prevention and diagnosis of a wide range of
 CC cancers. The present sequence represents the human R2 protein which is
 CC used in an example from the present invention

XX Sequence 389 AA;

Alignment Scores:

Pred. No.: 1,98e-165 Length: 389
 Score: 1454.00 Matches: 269
 Percent Similarity: 94.08% Conservative: 33
 Best Local Similarity: 83.80% Mismatches: 19
 Query Match: 78.76% Indels: 0
 DB: Gaps: 0

US-10-698-228-2 (1-1053) x AAB69051 (1-389)

QY 91 GAAGAGCCACTCTAGAAAGAGTTCGCGCGTTTGTTCATCTTCCATCCAGTACCCT 150
 DB 69 AspGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGluTyrHis 88
 QY 151 GTATTGTGAAATGPTAAACAGGCACAGGCTTCTTCGGACACAGAGAGGTGCGAC 210
 DB 89 AspIleTrpGlnMetTyrLysIleValGluAlaSerPheTrpThrAlaGluGluValAsp 108
 QY 211 TTATCAAGGATCTCCCTCACTGGAAACAAGCTTAAACAGACATGAGAGTACTTCTCTCT 270
 DB 109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128
 QY 271 CACATCTTAGCCTTTTTCAGCCAGTGATGGAATTTGTAATGAAATTTGGTGGAGCGC 330
 DB 129 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148
 QY 331 TTATGTCAGAGGTGCGAGTTCAGAGGTCGCTGTTCTTCATGGCTTCAAAATCTCATC 390
 DB 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
 QY 391 GAGATGTTCACTCAGAGATGTACAGTTGCTGATAGACACTTACATCAGAGATCCCAAG 450
 DB 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
 QY 451 AAAAGGGAATTTTATTAATCAATGAAACCATGCCCTATGTTTAAGAAAAAAGCAGAT 510
 DB 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysLysAlaAsp 208
 QY 511 TGGGCTTGGATGATAGATAGATAAATCTACTTTGGGAAAGAGTGGTGGCTTT 570
 DB 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaLaphe 228
 QY 571 GCTGCTGTAGAAGGAGTTTCTTCTCAGGATCTTTGCTGCTATATTCGGCTAAAGAAG 630
 DB 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 248
 QY 631 AGAGTCTTATCCAGGACTCACTTTTCCAAATGAACTCATCAGCAGAGATGAAGGACTT 690
 DB 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
 QY 691 CACTGTGACTTTGCTGCTGATGTCCAACTTAGTAATAAGCCTTCAGAGAAAGG 750
 DB 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
 QY 751 GTCAGGAGATCATTTGTCATGCTGTCAAAATTTGACGAGGAGTTTAAACAGAGCCTTG 810
 DB 289 ValArgGluIleIleIleAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 308
 QY 811 CCAGTTGGCCTCATTTGGAATGATTTGCAATTTTGTATGAAACAGTACATTCAGTTTGTAGCT 870
 DB 309 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 328

QY 871 GACAGATTACTGTGGAACCTTGGATTCTCAAGGTTTTTCAGGCAGAAAAATCCTTTTGAT 930
 DB 329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348
 QY 931 TTATGGAATAACATTTCTTTAGAGGAAAAACAATTTCTTTTGAGAAACGAGTTTCAGAG 990
 DB 349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368
 QY 991 TATCAGCGTTTTCAGATTATGCGAGAAACACACATACAGTCTTCCACTTGGATCGAGAT 1050
 DB 369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
 QY 1051 TTT 1053
 DB 389 Phe 389
 RESULT 5
 ABU07433
 ID ABU07433 standard; protein; 389 AA.
 AC ABU07433;
 XX 28-JAN-2003 (first entry)
 DT Protein differentially regulated in prostate cancer #36.
 XX Prostate cancer; gene expression; differential regulation;
 KW molecular marker; drug target; cancer detection; cancer diagnosis;
 KW cancer staging; cancer grading; cancer assessing; cancer monitoring.
 XX Homo sapiens.
 XX W0200281638-A2.
 PN 17-OCT-2002.
 XX 08-APR-2002; 2002WO-US010824.
 PF 06-APR-2001; 2001US-0281731P.
 PR 06-APR-2001; 2001US-0281732P.
 XX (ORIG-) ORIGENE TECHNOLOGIES INC.
 PA Sun Z, Jay G;
 PI WPI; 2003-058520/05.
 DR N-PSDB; ABX10335.
 DR Novel genes which are differentially regulated in prostate cancer, useful
 PT for diagnosing prostate cancer in prostate tissue sample and assessing
 PT therapeutic or preventive intervention in prostate cancer patients.
 XX Claim 1; Page 273-274; 416pp; English.
 PS The invention describes genes (I) which are differentially regulated in
 CC prostate cancer. (I) is useful for diagnosing a prostate cancer in a
 CC sample comprising prostate tissue, which involves determining the number
 CC of target genes which are differentially-regulated in the sample, where
 CC the number is indicative of the probability that the sample comprises
 CC prostate cancer. (I) is useful for assessing a therapeutic or preventive
 CC intervention in a subject having a prostate cancer, which involves
 CC determining the expression levels in a sample comprising prostate tissue
 CC of target genes which are differentially-regulated in prostate cancer.
 CC Preferably, the expression levels of at least 10 genes are determined.
 CC (I) is also useful for identifying agents that modulate a biological
 CC activity of a polypeptide differentially-regulated in prostate cancer
 CC cells, which involves contacting a polypeptide differentially-regulated
 CC in prostate cancer cells with a test agent under conditions effective for
 CC the test agent to modulate a biological activity of the polypeptide, and
 CC determining whether the test agent modulates the biological activity. (I)
 CC is useful as molecular markers, as drug targets, and for detecting,
 CC diagnosing, staging, grading, assessing, monitoring, prognosticating,
 CC preventing or treating, determining predisposition to diseases and

CC conditions especially relating to prostate cancer. (I) and its expression
 CC products are used in the diagnostic test to assay for presence of cancer
 CC e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in
 CC blood etc. (I) is useful for assessing cancer e.g., to determine the type
 CC of cancer, its stage of development, the nature of genetic defect, etc.
 CC The polypeptide encoded by (I) can be used as target for therapy or drug
 CC discovery. (I) can also be used for expressing the polypeptide and thus
 CC for searching specific binding partners of the polypeptide. (I) is useful
 CC in therapeutic applications to treat prostate cancer. The identification
 CC of specific genes, and groups of genes, expressed in pathways
 CC physiologically relevant to prostate cancer permits the definition of
 CC functional and disease pathways and the delineation of targets in these
 CC pathways which are useful in diagnostic, therapeutic, and clinical
 CC applications. This is the amino acid sequence of a protein differentially
 CC regulated in prostate cancer

SQ Sequence 389 AA;

Alignment Scores:

Pred. No.: 1-98e-165 Length: 389
 Score: 1454.00 Matches: 269
 Percent Similarity: 94.08% Conservatives: 33
 Best Local Similarity: 83.80% Mismatches: 19
 Query Match: 78.76% Indels: 0
 DB: 6 Gaps: 0

US-10-698-228-2 (1-1053) x ABU07433 (1-389)

QY 91 GAAGAGCCCTCTCTAAGAAAGATCTCGCGGTTGTGTCATCTTCCATTCAGTACCCCT 150
 DB 69 AspGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGluTyrHis 88
 QY 151 GATATTGGAAATGATATAACAGGCACAGGCTTCTTGGACACAGAGAGTGGAC 210
 DB 89 AspIleTrpGlnMetTyrIleValAlaGluAlaSerPheTrpThrAlaGluGluValAla 108
 QY 211 TTATCAAGGATCTCCCTCACTGGAAACAGCTTAAAGCAGATCAGAGTACTTCACTCT 270
 DB 109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128
 QY 271 CACATCTAGCCCTTTTTCAGCCAGTGTGAATGTAAATGAAATTTGGTGGAGCGC 330
 DB 129 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148
 QY 331 TTATGTCAGAGGTGTCAGGCTTCAGAGGCTCGCTTTCTATGGCTTTCAAATCTCATC 390
 DB 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
 QY 391 GAGATGTTTCACTCAGAGATGTACAGTTTGTGATAGACACTTACATCAGAGATCCCAAG 450
 DB 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
 QY 451 AAAAGGGAAATTTTATTAATGCAATTTGAAACCATGCCCTATGTTAAGAAAAAGCAGAT 510
 DB 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValIleLysLysAlaAsp 208
 QY 511 TGGGCTTCGGATGATACAGATGAAATCTACTTTTGGGGAAGAGTGGGCTTT 570
 DB 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValValAlaPhe 228
 QY 571 GCTGCTGTAGAAGGATTTTCTTCAGGATCTTTTGGCTATATTTCTGGCTAAAGAG 630
 DB 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 248
 QY 631 AGAGGTCTTATGCCAGGATCTCACTTTTCCATGAACTCATCAGCAGAGATGAAGACTT 690
 DB 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
 QY 691 CACTGTGACTTGTCTGTGATGTTTCCAAATCTAGTAAATAGCCTTCAGAAAGAG 750
 DB 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
 QY 751 GTCAGGGAGATCATTTGTCATGCTGTCAAATTTGACAGAGTCTTTTACAGAGGCTT 810

DB 289 ValArgGluIleIleIleAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 308
 QY 811 CCAGTTGGCTCTCATGGAATGCAATTTGATGAACAGTACATTCAGTTGTAGCT 870
 DB 309 ProValLysLeuGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 328
 QY 871 GACAGATTACTTGTGAACCTTGATTTCTCAAGGTTTTTCAGGCAGAAAAATCTTTTGTAT 930
 DB 329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348
 QY 931 TTATGGAACAACTTCTTTAGAGGAAAAACAAATTTCTTTGAGAAACGAGTTTCAGAG 990
 DB 349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPheGluLysArgValGlyGlu 368
 QY 991 TATCAGGCTTTTCAGTGTATGGCAGAAACACACAGATAACGTTCTTCACCTTGGATGCAGAT 1050
 DB 369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
 QY 1051 TTT 1053
 DB 389 Phe 389
 RESULT 6
 ADJ68753
 ID ADJ68753 standard; protein; 389 AA.
 AC ADJ68753;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE Human heat mitochondrial protein as a therapeutic target SeqID559.
 XX
 KW mitochondrial; human; screening assay; diabetes mellitus;
 KW Huntington's disease; osteoarthritis;
 KW Leber's hereditary optic neuropathy; LHON;
 KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
 KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
 KW neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;
 KW osteopathic; ophthalmological; cytostatic.
 XX
 OS Homo sapiens.
 XX
 PN WO2003087768-A2.
 XX
 PD 23-OCT-2003.
 XX
 PF 04-APR-2003; 2003WO-US010870.
 XX
 PR 12-APR-2002; 2002US-0372843P.
 PR 17-JUN-2002; 2002US-0389987P.
 PR 20-SEP-2002; 2002US-0412418P.
 XX
 PA (MITO-) MITOKOR.
 PA (BUCK-) BUCK INST AGE RES.
 XX
 PI Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;
 PI Warnock DE;
 DR WPI; 2003-845369/78.
 XX
 PT Identifying a mitochondrial target for drug screening assays and for
 PT treating diseases associated with altered mitochondrial function,
 PT comprises detecting a modified polypeptide in a sample and correlating
 PT with the disease.
 XX
 PS Claim 1; SEQ ID NO 559; 180pp; English.
 XX
 CC This invention relates to novel mitochondrial targets that can be used
 CC for therapeutic intervention in treating a disease associated with
 CC altered mitochondrial function. Specifically, it refers to a method for
 CC identifying proteins of the human heart mitochondrial proteome that are
 CC useful for drug screening assays, as well as therapeutic targets. The


```
QY 91 GAAGAGCCACTCTAAGAAAGAGTTCTCGCGGTTTGTTCATCTTTCCAAATCCAGTACCCT 150
DB 69 AspGluProLeuLeuArgGluAsnProArgArgPheValIlePheProIleGluTyrHis 88
QY 151 GATATTGGAAATGTATTAACAGGCACAGGCTTCTCTGAGCAGAGAGAGGTCGAC 210
DB 89 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 108
QY 211 TTATCAAGGATCTCCCTCACTGGAACAAGCTTAAGCAGATGAGAAGTACTCTCTCT 270
DB 109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128
QY 271 CACATCTTAGCCCTTTTTCAGCCAGTGTGAATTTGTAATGAAAAATTTGGTGGAGCGC 330
DB 129 HisValLeuAlaPhePheAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148
QY 331 TTATGTCAGAGGTGAGGTCCTCAGAGGCTCGCTTCTATGGCTTTCAAATTTCTCATC 390
DB 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
QY 391 GAGATGTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAGATCCCAAG 450
DB 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
QY 451 AAAAGGGAATTTTATTAATGCAATTAAGCAATCCCTATGTTAAAGAAAAAGCAGAT 510
DB 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 208
QY 511 TGGGCTTCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 570
DB 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaPhe 228
QY 571 GCTGCTGATAGAGAGTTTCTTCAGGATCTTTGCTGCTATATCTGCTTAAGAGAG 630
DB 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 248
QY 631 AGAGTCTTATGCCAGGACTCTTTTCCATGACTCATCAGCAGATGAGACATTT 690
DB 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
QY 691 CACTGTGACTTTGCTTGTGCTGATGTTCCAAATACTTAGTAAATAAGCCTTCAGAGAAAG 750
DB 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
QY 751 GTCAGGAGATCATTTGATGCTGCTCAAAATGTGACAGAGTTTAAACAGAGCCTTG 810
DB 289 ValArgGluIleIleAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 308
QY 811 CCAGTTGGCCTCATTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 870
DB 309 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 328
QY 871 GACAGATTACTTGTGGAATCTGATTTCTCAAAAGTTTTCAGGCAGAAAAATCTTTTGT 930
DB 329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348
QY 931 TTATGGAACAACTTTCTTAGAAGGAAAAACAAATTTCTTTGAGAAACGAGTTTCAGAG 990
DB 349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368
QY 991 TATCAGCCTTTTGGAGTTATGTCAGAAACACAGATTAAGCTTTCACCTTGGATGAGAT 1050
DB 369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
QY 1051 TTT 1053
DB 389 Phe 389
```

RESULT 9
ADN03789
ID ADN03789 standard; protein; 389 AA.
XX

```
AC ADN03789;  
XX 01-JUL-2004 (first entry)  
XX Antipsoriatic protein sequence #91.  
XX antipsoriatic; gene therapy; psoriasis; diagnosis.  
XX Homo sapiens.  
XX WO2004028479-A2.  
XX 08-APR-2004.  
XX 25-SEP-2003; 2003WO-US030907.  
XX 25-SEP-2002; 2002US-0414006P.  
XX (GETH ) GENENTECH INC.  
XX Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;  
XX Wu TD;  
XX WPI; 2004-305105/28.  
XX N-PSDB; ADN03788.  
XX New PRO nucleic acid or polypeptide, useful for preparing a  
XX pharmaceutical composition for diagnosing or treating psoriasis in a  
XX mammal.  
XX Claim 9; SEQ ID NO 183; 3069pp; English.  
XX The invention relates to novel polynucleotide and polypeptides for  
XX treating psoriasis or a sequence having at least 80% identity to the  
XX above sequences. The nucleic acid is useful for preparing a composition  
XX for diagnosing or treating psoriasis in a mammal. This sequence  
XX corresponds to one of the polypeptides of the invention.  
XX  
SQ Sequence 389 AA;  
Alignment Scores:  
Pred. No.: 1,98e-165 Length: 389  
Score: 1454.00 Matches: 269  
Percent Similarity: 94.08% Conservatives: 33  
Best Local Similarity: 83.80% Mismatches: 19  
Query Match: 78.76% Indels: 0  
DB: 8 Gaps: 0  
US-10-698-228-2 (1-1053) x ADN03789 (1-389)  
QY 91 GAAGAGCCACTCTAAGAAAGAGTTCTCGCGGTTTGTTCATCTTTCCAAATCCAGTACCCT 150  
DB 69 AspGluProLeuLeuArgGluAsnProArgArgPheValIlePheProIleGluTyrHis 88  
QY 151 GATATTGGAAATGTATTAACAGGCACAGGCTTCTCTGAGCAGAGAGAGGTCGAC 210  
DB 89 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 108  
QY 211 TTATCAAGGATCTCCCTCACTGGAACAAGCTTAAGCAGATGAGAAGTACTCTCTCT 270  
DB 109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128  
QY 271 CACATCTTAGCCCTTTTTCAGCCAGTGTGAATTTGTAATGAAAAATTTGGTGGAGCGC 330  
DB 129 HisValLeuAlaPhePheAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148  
QY 331 TTATGTCAGAGGTGAGGTCCTCAGAGGCTCGCTTCTATGGCTTTCAAATTTCTCATC 390  
DB 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168  
QY 391 GAGATGTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAGATCCCAAG 450  
DB 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
```

QY 451 AAAAGGGAATTTTATTAATGCAATGAAACCATGCCCTATGTTAAAGAAAAAGCAGAT 510
 ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
 Db GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 208
 QY 511 TGGGCTTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 570
 Db TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaPhe 228
 QY 571 GCTGCTGTAGAAGAGTCTTCTCAGGATCTTTTCTGCTATATCTTGGCTAAGAG 630
 Db AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLys 248
 QY 631 AGAGTCTTATCCAGGACTCACTTTTCCAAATGAATCATCAGCAGAGATCAAGGACT 690
 Db ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
 QY 691 CACTGTGACTTTGCTGCTGATGTTCCAAATCTAGTAAATAGCCTTCAGAAGAAAG 750
 Db HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
 QY 751 GTCAGGAGATCATGTTGATGCTGCTCAAAATGAGCAGAGTCTTTTACAGAGCCTTG 810
 Db ValArgGluIleIleAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 308
 QY 811 CCAGTTGGCTCATTTGGAAATGAATGCAATTTTGATGAAACAGTACATTCAGTTGTAGCT 870
 Db ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 328
 QY 871 GACAGATTACTTGTGGAACTTGATTTCTCAAAAGTCTTTCAGCAGAAATCTTTTGAT 930
 Db AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348
 QY 931 TTTATGAAAAATTTCTTTAGAAGAAAAACAAATTTCTTTGAGAAACAGTTTCAGAG 990
 Db PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368
 QY 991 TATCAGCTTTTGGAGTTATGTCAGAAACACAGATAAGCTTTCACCTTGGATGCAGAT 1050
 Db TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
 QY 1051 TTT 1053
 III
 Db 389 Phe 389
 RESULT 10
 ADN04444
 ID ADN04444 standard; protein; 389 AA.
 XX AC ADN04444;
 AC ADN04444;
 DT 01-JUL-2004 (first entry)
 XX DE Antipsoriatic protein sequence #415.
 XX KW antipsoriatic; gene therapy; psoriasis; diagnosis.
 XX OS Homo sapiens.
 XX PN WO2004028479-A2.
 XX PD 08-APR-2004.
 XX PF 25-SEP-2003; 2003WO-US030907.
 XX PR 25-SEP-2002; 2002US-0414006P.
 XX PA (GETH) GENENTECH INC.
 XX PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
 XX PI Wu TD;
 XX WPI; 2004-305105/28.

DR N-PSDB; ADN04443.
 XX New PRO nucleic acid or polypeptide, useful for preparing a
 PT pharmaceutical composition for diagnosing or treating psoriasis in a
 PT mammal.
 PS Claim 9; SEQ ID NO 838; 3069pp; English.
 XX The invention relates to novel polynucleotide and polypeptides for
 CC treating psoriasis or a sequence having at least 80% identity to the
 CC above sequences. The nucleic acid is useful for preparing a composition
 CC for diagnosing or treating psoriasis in a mammal. This sequence
 CC corresponds to one of the polypeptides of the invention.
 SX Sequence 389 AA;
 Alignment Scores:
 Pred. No.: 1,98e-165 Length: 389
 Score: 1454.00 Matches: 269
 Percent Similarity: 94.08% Conservative: 33
 Best Local Similarity: 83.80% Mismatches: 19
 Query Match: 78.76% Indels: 0
 DB: Gaps: 0
 US-10-698-228-2 (1-1053) x ADN04444 (1-389)
 QY 91 GAAGGCCACTCTTAAGAAAGATCTCGCGGTTTCTCATCTTTTCAATCCAGTACCCT 150
 Db 69 AspGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGluTyrHis 88
 QY 151 GATATTGGAAAAATGTATAAACAGGCACAGCTTCTCTGGACAGCAGAGAGGTGAC 210
 Db 89 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluValAsp 108
 QY 211 TTATCAAGATCTCCCTCACTCGAACACAGCTTAACAGCAGATGAGAGTACTCTCT 270
 Db 109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128
 QY 271 CACATCTTAGCCTTTTTCAGCCAGTGTGGAATTCGTAATGAAATTTCTGGAGCGC 330
 Db 129 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148
 QY 331 TTTAGTCAGGAGTGCAGGTTCCAGAGCTCGCTGTTTCTATGCTTTTCAAAATCTCATC 390
 Db 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
 QY 391 GAGAAATGTTCACTCAGAGATGTACAGTTGCTGATAGACACTTACATCAGAGATCCCAG 450
 Db 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
 QY 451 AAAAGGGAATTTTATTAATGCAATGAAACCATGCCCTATGTTAAAGAAAAAGCAGAT 510
 Db 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 208
 QY 511 TGGGCTTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 570
 Db 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaPhe 228
 QY 571 GCTGCTGTAGAAGAGTCTTCTCAGGATCTTTTCTGCTATATCTTGGCTAAGAG 630
 Db 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLys 248
 QY 631 AGAGTCTTATCCAGGACTCACTTTTCCAAATGAATCATCAGCAGAGATCAAGGACT 690
 Db 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
 QY 691 CACTGTGACTTTGCTGCTGATGTTCCAAATCTAGTAAATAGCCTTCAGAAGAAAG 750
 Db 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
 QY 751 GTCAGGAGATCATGTTGATGCTGCTCAAAATGAGCAGAGTCTTTTACAGAGCCTTG 810
 Db 289 ValArgGluIleIleAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 308

QY 931 TTTATGGAACATTTCTTTAGAGGAAACAAATTTCTTTCAGAAACGAGTTTCAGAG 990
Dd |||||
349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPheGluLysArgValGlyGlu 368
QY 991 TATCAGCGTTTGCAGTTATGCGAGAAACACAGATAACGCTTTCACCTGGATGAGAT 1050
Dd |||||
369 TyrGlnArgMetGlyValMetSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
QY 1051 TTT 1053
Dd |||
389 Phe 389
RESULT 12
ADQ09272
ID ADQ09272 standard; protein; 389 AA.
XX
AC ADQ09272;
XX
DT 23-SEP-2004 (first entry)
XX
DE Human RRM2 protein SEQ ID NO:457.
XX
KW thanatos-associated protein; THAP; THAP responsive gene; THAP family;
KW THAP responsive element; angiogenesis; inflammation; metastasis; cancer;
KW apoptosis; cardiovascular disease; neurodegenerative disease; chemokine;
KW antiangiogenic; antiinflammatory; cardiovascular; cytostatic;
KW neuroprotective; osteopathic; THAP modulator; THAP synthesis modulator;
KW human.
XX
OS Homo sapiens.
XX
PN WO2004055050-A2.
XX
PD 01-JUL-2004.
XX
PF 10-DEC-2003; 2003WO-IB006434.
XX
PR 10-DEC-2002; 2002US-0432699P.
PR 03-JUN-2003; 2003US-0485027P.
XX
PA (ENDO-) ENDOCUBE SAS.
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Girard J, Amalric F, Roussigne M, Clouaire T;
DR WPI; 2004-525034/50.
DR N-PSDB; ADQ09273.
XX
PT Modulating expression of a Thanatos (death)-Associated Protein (THAP)
PT responsive gene for preventing or treating e.g. cancer or inflammation,
PT comprises modulating the interaction of a THAP polypeptide with a nucleic
PT acid.
XX
PS Example 47; SEQ ID NO 457; 612pp; English.
XX
CC The present invention describes a method for modulating the expression of
CC a thanatos (death)-associated protein (THAP) responsive gene. The method
CC comprises modulating the interaction of a THAP-family polypeptide or its
CC biological fragment with a nucleic acid, and so enhancing or repressing
CC the expression of the THAP responsive gene. Also described: (1) a method
CC of modulating the expression of a gene responsive to a THAP/chemokine
CC complex; (2) a pharmaceutical composition comprising a THAP responsive
CC element in a pharmaceutical carrier; (3) a transcription factor decoy
CC consisting essentially of a THAP responsive element; (4) a cell
CC comprising a transcription factor decoy described above; (5) methods of
CC modulating the interaction between a nucleic acid and a THAP-family
CC polypeptide or its biological fragment, or a nucleic acid and a
CC THAP/chemokine complex; (6) a vector packaging cell line comprising a
CC cell comprising a viral vector which comprises a promoter operably linked
CC to a nucleic acid encoding a THAP-family polypeptide or its biological
CC fragment; (7) a method of constructing a cell which expresses a
CC recombinant THAP-family polypeptide; (8) a method of ameliorating
CC symptoms associated with a condition mediated by a THAP/chemokine complex

CC ; (9) methods of identifying a test compound that modulates transcription
CC at a THAP responsive element or that modulates the transport of a
CC chemokine into the nucleus; (10) methods for reducing the symptoms
CC associated with a condition selected from excessive or insufficient
CC angiogenesis, inflammation, metastasis of a cancerous tissue, excessive
CC or insufficient apoptosis, cardiovascular disease and neurodegenerative
CC diseases; symptoms associated with a condition resulting from the
CC activity of a chemokine or a THAP-family polypeptide in an individual; or
CC symptoms associated with transcriptional repression or activation
CC mediated by a THAP-family polypeptide in an individual; (11) a vector
CC comprising a THAP responsive promoter operably linked to a nucleic acid
CC encoding a detectable product; (12) a genetically engineered cell
CC comprising the vector described above or that expresses a THAP-family
CC polypeptide or its biological fragment; (13) an in vitro transcription
CC reaction comprising a nucleic acid comprising a THAP responsive promoter,
CC ribonucleotides and an RNA polymerase; and (14) an isolated mutant THAP-
CC family polypeptide that does not bind to a chemokine. The pharmaceutical
CC composition has antiangiogenic, antiinflammatory, cardiovascular,
CC cytostatic, neuroprotective and osteopathic activities, and can be used
CC as a THAP and THAP synthesis modulator. The composition can be used for
CC modulating the expression of a THAP responsive gene. Modulation is useful
CC for reducing symptoms of conditions such as excessive or insufficient
CC angiogenesis, inflammation, metastasis of a cancerous tissue, excessive
CC or insufficient apoptosis, cardiovascular disease or neurodegenerative
CC diseases. The present sequence is used in the exemplification of the
CC present invention.
XX
SQ Sequence 389 AA;

Alignment Scores:

Pred. No.:	1-98e-165	Length:	389
Score:	1454.00	Matches:	269
Percent Similarity:	94.08%	Conservative:	33
Best Local Similarity:	83.80%	Mismatches:	19
Query Match:	8	Indels:	0
DB:		Gaps:	0

US-10-698-228-2 (1-1053) x ADQ09272 (1-389)

QY	91	GAAGAGCCACTCTCTAAGAAAGAGTTCTCGCGGTTTGTCATCTTTCCAAATCCAGTACCCT	150
Dd	69	AspGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGluThrHis	88
QY	151	GATATTTGGAATGTATAACAGGACAGGCTTCCTCTGGACAGCAGAGAGGTGCAC	210
Dd	89	AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp	108
QY	211	TTATCAAGGATCTCCTCACTGGAACAGCTTAAAGCAGATGAGAGTACTTCTCTCT	270
Dd	109	LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer	128
QY	271	CACATCTTAGCCCTTTTTCAGCCAGTGATGGAATTTGAAATTTGAAATTTGGTGAGCGC	330
Dd	129	HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg	148
QY	331	TTTATGTCAGGAGTGCAGGTTCCAGAGGCTCGTGTCTTCTATGGCTTCAATTTCTCATC	390
Dd	149	PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet	168
QY	391	GAGAAATGTTCCACTCAGAGATGTACAGTTTCTCATGACACTTACATCAGAGATCCCAAG	450
Dd	169	GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys	188
QY	451	AAAAGGAAATTTTATTTAATGCAATTTGAACCATGCCCTATGTATTAAGAAAAAGCAGAT	510
Dd	189	GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp	208
QY	511	TGGGCTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	570
Dd	209	TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaAlaPhe	228
QY	571	GCTGCTGTAGAAGAGTTTCTTCTCAGGATCTTTCTGCTGCTATATTCTGCTAAAGAG	630

Db 229 AlaAlaValGluGlyLeuPhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 248
QY 631 AGAGGTCCTATGCCAGGACTCACTTTTCCAAATGAACCTCATCAGCAGAGATGAAGACTT 690
Db 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
QY 691 CACTGTGACTTTGCTTGCCTGATGTTTCCAAATCTAGTAAATGAAGCCTTCAGAGAAAGG 750
Db 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
QY 751 GTCAGGAGATCATTTGTCATGCTGCAAAATGACGAGAGTTTAAACAGAGCCTTG 810
Db 289 ValArgGluIleIleAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 308
QY 811 CCAGTTGGCTCATTTGGAATGAATTCATTTGATGAAACAGTACATTCAGTTGTAGCT 870
Db 309 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTrpIleGluPheValAla 328
QY 871 GACAGATTACTTGTGGAATCTGGATCTCAAAAGTTTTCAGCAGAAATCTCTTTGAT 930
Db 329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348
QY 931 TTTATGGAACATTTCTTTAGAGAAACAAATTTCTTTGAGAAACGAGTTTCAGAG 990
Db 349 PheMetGluAsnIleSerLeuGlyLysThrAsnPheGluLysArgValGlyGlu 368
QY 991 TATCAGCTTTTGCAGTTATGTCAGAAACACAGATAAGCTTTCACCTTGATGAGAT 1050
Db 369 TyrGlnArgMetGlyValMetSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
QY 1051 TTT 1053
Db 389 Phe 389
RESULT 13
ABM80258
ID ABM80258 standard; protein; 389 AA.
XX
AC ABM80258;
XX
DT 18-NOV-2004 (first entry)
XX
DE Tumour-associated antigenic target (TAT) polypeptide PRO62893, SEQ:652.
XX
KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW tumour; diagnosis; cell proliferative disorder; breast cancer;
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW central nervous system cancer; bladder cancer; pancreatic cancer;
KW cervical cancer; melanoma; leukaemia; hybridisation probe;
KW chromosome identification; chromosome mapping; gene mapping;
KW gene therapy; cytostatic.
XX
OS Homo sapiens.
XX
PN WO2004030615-A2.
XX
PD 15-APR-2004.
XX
PF 29-SEP-2003; 2003WO-US028547.
XX
PR 02-OCT-2002; 2002US-0414971P.
XX
PA (GETH) GENENTECH INC.
XX
PI Wu TD, Zhang Z, Zhou Y;
XX
DR WPI; 2004-347921/32.
DR N-PSDB; ACN37637.
XX
PT New tumor-associated antigenic target polypeptides and nucleic acids,
PT useful in preparing a medicament for treating or detecting a
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT prostate cancer or tumor.

XX Claim 12; SEQ ID NO 652; 7273pp; English.
PS
XX
CC The invention relates to human tumour-associated antigenic target (TAT)
CC polypeptides, and their related nucleic acids. The TAT polypeptides are
CC overexpressed in cancer tissues compared to normal tissues, and may thus
CC serve as effective targets for the diagnosis and treatment of cancer in
CC mammals. The invention also relates to nucleic acid and polypeptide
CC sequences at least 80% identical to the TAT nucleic acids and
CC polypeptides; expression vectors and host cells comprising a TAT nucleic
CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC TAT polypeptide; and methods and compositions for the treatment or
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC antibodies, antagonists, binding molecules and compositions are useful
CC for diagnosing or treating a cell proliferative disorder associated with
CC increased TAT expression, particularly cancers such as breast cancer,
CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC cancer, pancreatic cancer, cervical cancer, cancers of the central
CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC used as hybridisation probes, in chromosome and gene mapping, in
CC chromosome identification and in gene therapy. The present sequence
CC represents a TAT polypeptide of the invention
XX
SQ Sequence 389 AA;
Alignment Scores:
Pred. No.: 1,98e-165 Length: 389
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: 8 Gaps: 0
US-10-698-228-2 (1-1053) x ABM80258 (1-389)
QY 91 GAAGAGCCACTCTTAAGAAAGAGTCTCTCGCGCTTCTCATCTTTCCAAATCCAGTACCCT 150
Db 69 AspGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGluTrpHis 88
QY 151 GATATTGGAAATATGATAACAGCAGGCTTCTCTGACAGAGAGAGAGTCCGAC 210
Db 89 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 108
QY 211 TTATCAAGATCTCCCTCCTCAGAACAACTTAAGCAGATGAGAGTACTCATCTCT 270
Db 109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128
QY 271 CACATCTTAGCCTTTTTCAGCAGCAGTGGATTCGTAATGAAATTTGGTGGAGCC 330
Db 129 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148
QY 331 TTTAGTCAGAGGTGCAGGTTCAGAGGCTCGCTGTTTCTATGGCTTTCAAATTTCTCATC 390
Db 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
QY 391 GAGAATGTTCACTCAGAGATGTACAGTTTCTGATAGACACTTACATCAGAGATCCCAAG 450
Db 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
QY 451 AAAAGGGAATTTTATTATGCAATGAAACCATGCCCTATGTTAGAAAAAGCAGAT 510
Db 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 208
QY 511 TGGGCTTCGATGGATAGCAGATAGAAATCTACTTTTGGGAAAGAGTGGTGGCTTTT 570
Db 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTrpGlyGluArgValAlaPhe 228
QY 571 GCTGCTGTAGAAGGAGTTTCTTCTCAGGATCTTTTCTGCTGTATATTTCTGGCTAAAGAG 630
Db 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 248
QY 631 AGAGGTCTTATGCCAGGACTCACTTTTCCAAATGAACTCATCAGCAGAGATGAAGACTT 690

Db 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
QY 691 CACTGTGACTTTGCTGCTGATGTTCCAACTACTAGTAATAAGCCCTCAGAGAAAG 750
Db 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
QY 751 GTCAGGGAGATCAATGTTGATGCTGTCAAAATTTGAGCAGGAGTTTTTAACAGAGCCCTG 810
Db 289 ValArgGluIleIleHisAsnAlaValArgLleGluGlnGluPheLeuThrGluAlaLeu 308
QY 811 CCAGTTGGCCTCATTTGGAATGAATTCATTTTGATGAAACAGTACATTGAGTTTGTAGCT 870
Db 309 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 328
QY 871 GACAGATTACTTGTGGAACTTGATCTCAAGGTTTTTTCAGGCAGAAATCCTTTTGAT 930
Db 329 AspArgLeuMetGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348
QY 931 TTTATGGAAAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACGAGTTTCAGAG 990
Db 349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPheGluLysArgValGlyGlu 368
QY 991 TATCAGCGTTTTCAGTTATGGCAGAAACCCAGATAACGTTCTACCTTGGATGCAGAT 1050
Db 369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
QY 1051 TTT 1053
Db 389 Phe 389
RESULT 14
AAB43902
ID AAB43902 standard; protein; 413 AA.
XX
AC AAB43902;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human cancer associated protein sequence SEQ ID NO:1347.
XX
KW Human; cancer associated gene; cancer antigen; detection; cancer;
KW diagnosis; cytostatic; proliferative; vulnery; immunomodulator;
KW antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
KW antinflammatory; antithyroid; antiallergic; antibacterial; cardiac;
KW dermatological; neuroprotective; thrombolytic; coagulant; neutropenic;
KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW allergic reaction; graft versus host disease; organ rejection;
KW haemostatic; thrombolytic; cardiovascular disorder; infection;
KW neurological disease; drug screening.
XX
OS Homo sapiens.
XX
PN WO200055350-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US005882.
XX
PR 12-MAR-1999; 99US-0124270P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2000-587533/55.
DR N-PSDB; AAC78111.
XX
PT Novel isolated nucleic acids comprising sequences encoding peptides
XX useful for treating or diagnosing e.g. cancer.
XX
PS Claim 11; Page 2000-2002; 2352pp; English.

XX
CC AAC77607 to AAC78448 encode the human cancer associated proteins given in
CC AAB43398 to AAB44239. The proteins can have activities based on the
CC tissues and cells the genes are expressed in. Example of activities
CC include: cytostatic; proliferative; vulnery; immunomodulator;
CC antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
CC antinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
CC dermatological; neuroprotective; cardiac; thrombolytic; coagulant;
CC neutropenic; vasotropic; antipsoriatic and antiangiogenic. The
CC polynucleotides and polypeptides can be used for preventing, treating or
CC ameliorating medical conditions and diagnosing pathological conditions.
CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC the present invention may be used to treat immune disorders by activating
CC or inhibiting the proliferation, differentiation or mobilisation of
CC immune cells, to treat disorders of haematopoietic cells, autoimmune
CC disorders, allergic reactions, graft versus host disease and organ
CC rejection, modulate haemostatic or thrombolytic activity, modulate
CC inflammation, cancers, cardiovascular disorders, neurological disease and
CC bacterial or viral infections. The peptides, nucleotides, antibodies,
CC agonists and antagonists may be also be used in drug screens. AAC78449 to
CC AAC78457 and AAB44240 represent sequences used in the exemplification of
CC the present invention
XX
SQ Sequence 413 AA;
Alignment Scores:
Pred. No.: 2,04e-165 Length: 413
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: 3 Gaps: 0
US-10-698-228-2 (1-1053) x AAB43902 (1-413)
QY 91 GAAGACCCACTCTTAAGAAAGAGTTCTCGCGGTTTGTCTATCTTTCCAATCCAGTACCCT 150
Db 93 AspGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGluTyrHis 112
QY 151 GATATTGGAAATGTATAAACAGGACAGGCTTCCTCTGGACAGACAGAGAGTCCAC 210
Db 113 AspIleTgPlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 132
QY 211 TTATCAAGAGATCTCCCTCACTCGAACAGCTTAAGCAGATGAGAGTACTTCATCTCT 270
Db 133 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgGlyPheIleSer 152
QY 271 CACATCTTAGCCCTTTTTCAGCCAGTGTGATGGAATTTGTAATGAAATTTTGGTGGAGCGC 330
Db 153 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 172
QY 331 TTTAGTCAGAGGTGCAGGTTCCAGAGCTCGCTGTTTCTATGCTTTCAATTTCTATC 390
Db 173 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 192
QY 391 GAGAAATGTTCACTCAGAGATGTACAGTTTCTGTGTAGACACTTACATCAGAGATCCCAAG 450
Db 193 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 212
QY 451 AAAAGGGAATTTTATTATTAATGCAATTAACCACTATGCTTATGTAAGAAAAAGCAGAT 510
Db 213 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 232
QY 511 TGGGCTTGGATGGATGAGATAGAGAAATCTACTTTTGGGGAAGAGTGGTGGCCCTTT 570
Db 233 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaAlaPhe 252
QY 571 GCTGCTGTAGAAGAGTTTCTTCTCAGGATCTTTTGTCTGTATATTCTGCTAAAGAG 630
Db 253 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 272
QY 631 AGAGGTCTTATGTCAGGACTCACCTTTTCCAACTCAATCAGCAGAGATGAAGACTT 690

Db 273 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 292

QY 691 CACTGTGACTTTGCTTGGCTGATGTTCCCAATCTAGTAATAAAGCTTCAGAGAAGG 750

Db 293 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 312

QY 751 GTCAGGAGATCATTTGCTGATGCTGCAAAATTTGACGACGAGTTTAAACAGAGCCTTG 810

Db 313 ValArgGluIleIleAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 332

QY 811 CCAGTTGGCTCATTTGGAATGATTCATTTTGATCAAAACAGTACATTGAGTTTGTAGCT 870

Db 333 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnIleGluPheValAla 352

QY 871 GACAGATTACTTGTGGAATCTGATCTCAAGGTTTTCAGCGAAGAAATCTTTTGAT 930

Db 353 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 372

QY 931 TTTATGGAAAAATTTCTTTAGAGGAAAAACAAATTTCTTTGAGAAACGAGTTTCAGAG 990

Db 373 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 392

QY 991 TATCAGCTTTGTCAGTTATGTCAGAAACACAGATAAGCTTTCACCTTGATGTCAGAT 1050

Db 393 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 412

QY 1051 TTT 1053

Db 413 Phe 413

RESULT 15

AAU28017

ID AAU28017 standard; protein; 453 AA.

AC AAU28017;

DT 18-DEC-2001 (first entry)

DE Human contig polypeptide sequence #170.

XX Mammal; human; rhesus monkey; baker's yeast; fission yeast; Norway rat; mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia; cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation; nervous system disorder; inflammatory disorder; cell differentiation; angiogenesis; stem cell growth factor; actinin; inhibin; cartilage; burn; genetic disorder; bone regeneration; tendon; ligament; tissue repair; cytostatic; antirheumatic; antiarthritic; vulnery; antiinflammatory; antibacterial; immunosuppressive; vasotropic; antiparkinsonian; neuroprotective; osteopathic; antidiabetic; antiallergic; immunostimulant; analgesic; gene therapy.

OS Homo sapiens.

OS Synthetic.

XX WO200164834-A2.

PN 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US004926.

PF 28-FEB-2000; 2000US-00515126.

PR 18-MAY-2000; 2000US-00577409.

PR 17-JUN-2000; 2000US-00597707.

PR 14-JUL-2000; 2000US-00616807.

PR 19-SEP-2000; 2000US-00664641.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;

PI Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;

PI Drmanac R;

XX WPI; 2001-589862/66.

DR N-PSDB; AAS44917.

XX Novel polypeptides and nucleic acids obtained from cDNA libraries prepared from various human tissues, for diagnosis, treatment of cancer, neurological, inflammatory disorders and for use in arrays for detection.

PT Claim 10; Page 146-147; 153pp; English.

XX Sequences AAU27676-AAU28019 represent full-length polypeptides and contig polypeptides of the invention. The proteins and their associated DNA sequences are useful for the treatment, diagnosis and prevention of various types of disorder in a mammalian subject such as a human, dog, monkey, mouse, hamster or rat. The disorders include cancers such as leukaemia, lymphoma and neuroblastoma, autoimmune disorders such as multiple sclerosis, connective tissue disease, rheumatoid arthritis, diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system disorders such as Parkinson's disease, Alzheimer's disease, Huntington's chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and Wernicke disease, inflammatory disorders such as nephritis, Crohn's disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory bowel disease. The sequences exhibit activity relating to angiogenesis, cell proliferation, cell differentiation, stem cell growth factor, cell activin or inhibin. Therefore, they can be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, accidental damage or genetic disorders. The sequences may also be used for regeneration of bone, cartilage, tendons and ligaments and in tissue repair and burn healing.

CC Note: Some sequences for this patent did not form part of the printed specification, but were obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 453 AA;

Alignment Scores:

Pred. No.:	2.13e-165	Length:	453
Score:	1454.00	Matches:	269
Percent Similarity:	94.08%	Conservative:	33
Best Local Similarity:	83.80%	Mismatches:	19
Query Match:	78.76%	Indels:	0
DB:	4	Gaps:	0

US-10-698-228-2 (1-1053) x AAU28017 (1-453)

QY 91 GAGAGCCACTCTTACGAAAGAGTTCTCGCGGTTTGTCTATCTTCCATCCAGTACCCT 150

Db 133 AspGluProLeuLeuArgGluAsnProArgArgPheValIlePheProIleGluTyrHis 152

QY 151 GATATTGGAAATGTATAACACGACGAGCTTCTCTCGACAGCAGAGAGTTCGAC 210

Db 153 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheThrAlaGluGluValArg 172

QY 211 TTATCAAGGATCTCCCTCCTCAGTGGACCAAGCTTAAAGCAGATGAGAGTACTCTCT 270

Db 173 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 192

QY 271 CACATCTAGCCTTTTTCAGCCAGTGGAAATGTAATGAAATTTGGTGGAGCGC 330

Db 193 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 212

QY 331 TTTTAGTCAGGAGTTCAGAGTTCCAGAGCTCGCTGTTTCTATGGCTTTCAAATCTCTCATC 390

Db 213 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 232

QY 391 GAGAATGTTTCACCTCAGAGATGTACAGTTGCTGTATAGACACTTACATCAGAGATCCCAG 450

Db 233 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 252

QY 451 AAAGGGAATTTTATTATTAATGCAATTGAAACCAATGCCCTATGTTTAAAGAAAACAGAT 510

Db 253 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 272

QY 511 TGGGCTTCGATGGATAGCAGATAGAAATCTACTTTTGGGAAAGAGTGTGCGCTTT 570

Db 273 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrglyGluArgValValAlaPhe 292
QY 571 GCTGCTGTAGAGGAGTTTCTTCTCAGGATCTTTTCTGCTATATCTTGGCTAAAGAAG 630
Db 293 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLys 312
QY 631 AGAGTCTTATGCCAGGACTCATTCTTCCAAATGAATCATCAGCAGAGATGAAGGACTT 690
Db 313 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 332
QY 691 CACTGTGACTTTGCTTGGCTGATGTTCCAATACTTAGTAATAAGCCTTCAGAAAGG 750
Db 333 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 352
QY 751 GTCAGGGAGATCATTTGATGCTGTCATAAATTGACGAGGAGTCTTTTAAACAGAGCCTTG 810
Db 353 ValArgGluIleIleGlnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 372
QY 811 CCAGTTGGCCTCATTTGGAATGAATTCATTTTGTATGAACACAGTACATTCAGTTTGTAGCT 870
Db 373 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrlleGluPheValAla 392
QY 871 GACAGATTACTTGTGGAACTTGGATCTCAAAAGGTTTTTCAGGCAGAAATCCTTTTGAT 930
Db 393 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 412
QY 931 TTTATGGAACAACTTTCTTAGAAGGAAACAAATTTCTTGAGAAACGAGTTTCAGAG 990
Db 413 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 432
QY 991 TATCAGCGTTTTCAGTATGTCAGAAACACAGATTAACGTCCTTCACCTTGGATGCAGAT 1050
Db 433 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 452
QY 1051 TTT 1053
Db 453 Phe 453

Search completed: October 30, 2005, 06:28:41
Job time : 79.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 30, 2005, 05:15:40 ; Search time 15.75 Seconds
(without alignments)
12865.556 Million cell updates/sec

Title: US-10-698-228-2

Perfect score: 1846

Sequence: 1 atggggccggaaagcc.....tcaccttgatgcagatttt 1053

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+_n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10698228/runat_26102005_100608_3995/app_query.fasta_1.2446
-DB=PIR_79 -QFMT=fastan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10698228@cgn_1_1_39@runat_26102005_100608_3995 -NCPU=6 -ICPU=3
-NO WMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PIR_79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1821	98.6	366	2 T46249	hypothetical prote
2	1454	78.8	389	2 S25854	ribonucleoside-dip
3	1443	78.2	390	1 S86735	ribonucleoside-dip
4	1421	77.0	386	2 S27153	ribonucleoside-dip
5	1331	72.1	384	2 S24585	ribonucleoside-dip
6	1280	69.3	319	2 T30782	ribonucleoside-dip
7	1280	69.3	319	2 B72154	E4L protein - vari
8	1279	69.3	319	1 RDVZV	ribonucleoside-dip
9	1279	69.3	333	2 T28466	ribonucleoside-dip
10	1266	68.6	333	2 H36839	C8L protein - vari
11	1258.5	68.2	348	2 T30470	ribonucleoside-dip
12	1217	65.9	299	1 RDSS2R	ribonucleoside-dip
13	1214	65.8	381	2 T18876	hypothetical prote
14	1169.5	63.4	399	1 A26916	ribonucleoside-dip

15	1161.5	62.9	391	2 T39992	ribonucleoside-dip
16	1161.5	62.9	391	2 S34808	ribonucleoside-dip
17	1128.5	61.1	329	2 T03688	ribonucleoside-dip
18	1070.5	58.0	340	2 S68538	ribonucleoside-dip
19	1024.5	55.5	349	2 B49412	ribonucleoside-dip
20	950	51.5	324	2 T17978	probable ribonucle
21	805	43.6	345	2 S59744	ribonucleoside-dip
22	681	36.9	668	2 T29884	hypothetical prote
23	638	34.6	327	1 RDVZAS	ribonucleoside-dip
24	385	20.9	331	2 A84389	ribonucleoside red
25	384.5	20.8	415	2 A83502	ribonucleoside red
26	358.5	19.4	346	2 G81728	ribonucleoside-dip
27	357.5	19.4	346	2 E71466	probable ribonucle
28	335	18.1	346	2 B72010	ribonucleoside-dip
29	335	18.1	346	2 F86613	ribonucleoside red
30	329	17.8	306	1 WMBE18	ribonucleoside-dip
31	324	17.6	303	2 T47526	ribonucleoside-dip
32	322	17.4	320	2 T42563	ribonucleoside-dip
33	318	17.2	321	1 WMBE11	ribonucleoside-dip
34	302.5	16.4	314	1 WMBE14	ribonucleoside-dip
35	301	16.3	302	1 WMBE12	ribonucleoside-dip
36	294.5	16.0	347	2 C87281	hypothetical prote
37	292.5	15.8	340	1 WMBE17	ribonucleoside-dip
38	290	15.7	305	2 S55655	ribonucleoside-dip
39	290	15.7	328	2 C97781	hypothetical prote
40	289.5	15.7	337	1 WMBE32	ribonucleoside-dip
41	288.5	15.6	337	1 WMBE12	ribonucleoside-dip
42	288.5	15.6	341	2 C71858	ribonucleoside-dip
43	286.5	15.5	341	2 D64565	ribonucleoside-dip
44	285	15.4	305	2 T03155	ribonucleoside-dip
45	285	15.4	324	2 B71655	ribonucleoside-dip

ALIGNMENTS

RESULT 1

T46249

hypothetical protein DKFZp761E1312.1 - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004

C;Accession: T46249

R;Ansonge, W.; Winkler, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, January 2000

A;Reference number: Z23028

A;Accession: T46249

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-366 <AAA>

A;Cross-references: UNIPROT:Q9NTD8; EMBL:AL137348

A;Experimental source: adult amygdala; clone DKFZp761E1312

C;Genetics:

A;Note: DKFZp761E1312.1

C;Superfamily: ribonucleoside reductase small subunit

Alignment Scores:

Pred. No.:	2,85e-157	Length:	366
Score:	1821.00	Matches:	351
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	98.65%	Indels:	0
DB:	2	Gaps:	0

US-10-698-228-2 (1-1053) x T46249 (1-366)

Qy	1	ATGGGCGACCCGAAAGCGCGCGCGCTGGATCAGGATGAGAGATCATCTTCA	60
Db	16	MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer	35
Qy	61	GACACCAACGAAAGTGAATAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTGC	120
Db	36	AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg	55
Qy	121	CGTTTGTTCATCTTTCATCCAGTACCCTGATATTGGAAATGTATTAACAGGCACAG	180

```
|||||
56 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 75
QY 181 GCTCTCTCTGACACAGAGCTCGACTTATCAAGAGATCTCCCTCACTGGACAG 240
Db 76 AlaSerPheTrpTrpAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 95
QY 241 CTTAAAGCAGATGAGAAGTACTCTATCTCTACATCTTAGCCCTTTTTCAGCCAGTGTAT 300
Db 96 LeuLysAlaAspGluLysTyrPheIleSerHisLeuAlaPhePheAlaAlaSerAsp 115
QY 301 GGAATTGTAATGAAATTTGTGTGAGCGCTTTAGTCAGAGGTCAGGTTCCAGAGGCT 360
Db 116 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 135
QY 361 CGCTGTTCTATGGCTTCAATTTCTCATCGAGATGTTCTCATGAGATGATACAGTTTG 420
Db 136 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 155
QY 421 CTGATAGACACTATACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 480
Db 156 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 175
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGCGCTTTCGATGGATAGCAGATAGAAA 540
Db 176 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 195
QY 541 TCTACTTTTGGGAAAGAGTGTGGCTTGTCTGTAGAGGAGTTTCTTCTCAGGA 600
Db 196 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 215
QY 601 TCTTTTGTCTATATTCTGGCTTAAAGAGAGAGGTCTTATCCAGGACTCACTTTTTC 660
Db 216 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 235
QY 661 AATGAACCTCATCAGCAGATGAAGGACTTCACTGTGACTTTTCTTGGCTGATGTTCCAA 720
Db 236 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 255
QY 721 TACTTAGTAAATAGCTTCAGAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 780
Db 256 TyrLeuValAsnLysProSerGluGluArgValArgGluIleValAlaValLys 275
QY 781 ATTGACGAGGAGCTTTTACAGAGCTTGCAGATTGGCTCATTTGGAATGATTCATT 840
Db 276 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 295
QY 841 TTGATGAAACAGTACATTGAGTTGTAGCTGACAGATTACTTGTGAACTTGATTTCTCA 900
Db 296 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 315
QY 901 AAGGTTTTTCAGGAGAAAATCCTTTTGTATTTATGGAATAACATTCTTTTAGAGAAA 960
Db 316 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluLys 335
QY 961 ACAAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGGTTATGCGCAAGACC 1020
Db 336 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 355
QY 1021 ACAGATAACGCTCTTCACTTGGATGATGAGATTTT 1053
Db 356 ThrAspAsnValPheThrLeuAspAlaAspPhe 366
```

RESULT 2

S25854
ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain - human
N:Alternate names: ribonucleotide reductase M2 chain; ribonucleotide reductase small chain
C:Species: Homo sapiens (man)
C:Date: 08-Jun-1994 #sequence revision 10-Nov-1995 #text change 09-Jul-2004
C:Accession: S25854
R:Pavloff, N.; Rivard, D.; Masson, S.; Shen, S.H.; Mes-Masson, A.M.
DNA Seq. 2, 227-234, 1992
A:Title: Sequence analysis of the large and small subunits of human ribonucleotide reductase

A:Reference number: S25853; MUID:92329977; PMID:1627826
A:Accession: S25854
A:Molecule type: mRNA
A:Residues: 1-389 <PAV>
A:Cross-references: UNIPROT:P31350; EMBL:X59618; MID:g36154; PIDN:CAA42181.1; PID:g36155
C:Genetics:
A:Gene: GDB:RRM2
A:Cross-references: GDB:120358; OMIM:180390
A:Map position: 2p25-2p24
C:Superfamily: ribonucleoside reductase small subunit
C:Keywords: deoxyribonucleoside biosynthesis; iron; metalloprotein; oxidoreductase
F:138,169,172,232,266,269 Binding site: 2Fe-O cluster (Asp, Glu, His, Glu, His) #status predicted
F:176/Active site: Tyr (stable tyrosyl radical) #status predicted

Alignment Scores:
Pred. No.: 6,58e-124 Length: 389
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
Db: 2 Gaps: 0

US-10-698-228-2 (1-1053) x S25854 (1-389)

```
QY 91 GAAGAGCCACTCTTAAGAAAGAGTTCTCGCGGTTTGTCTATCTTCCAATCCAGTACCCT 150
Db 69 AspGluProLeuLeuArgGluAsnProArgArgPheValIlePheProIleGluTyrHis 88
QY 151 GATATTTGAAATGATATAACAGGACAGGCTTCTCTGGACAGAGAGAGGTGCAAC 210
Db 89 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 108
QY 211 TTATCAAGAGATCTCCCTCACTCGAACACAGCTTAAACACAGATGAGAGTACTTCTCTCT 270
Db 109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128
QY 271 CACATCTTAGCCCTTTTTCAGCCAGTCATGGAATTTAAATGAAATTTTGGTGGAGCGC 330
Db 129 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148
QY 331 TTATGTCAGAGGTGAGGTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAAATCTCATC 390
Db 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
QY 391 GAGAAATGTTCACTCAGAGATGTACAGTTTCTCTATAGACACTTACATCAGAGATCCCAAG 450
Db 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
QY 451 AAAAGGGAATTTTATTATGCAATTTGAAACCATGCCCTATGTTAAAGAAAAAGCAGAT 510
Db 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 208
QY 511 TGGGCTTTCGATGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 570
Db 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaAlaPhe 228
QY 571 GCTGCTGTAGAGAGGTTTCTTCTCAGAGTCTTTTCTGCTGCTATATTTCTGCTGAAGNAG 630
Db 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 248
QY 631 AGAGGTTCTTATGCCAGGACTCAGCTTTTCCAAATGAACTCATCAGCAGATGAAAGGACTT 690
Db 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
QY 691 CACTGTGACTTTTCTTCTGCTGATGTTTCCAAATGAACTCATCAGCAGATGAAAGGACTT 750
Db 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
QY 751 GTCAGGAGATCATTTGATGCTGCAAAATTTAGCAGAGGAGTTTAAACAGAGCCCTTG 810
Db 289 ValArgGluIleIleIleAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 308
QY 811 CCAGTTGGCCTCATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 870
```

```

Db      309 ProValLeuLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 328
QY      871 GACAGATTACTTGTGAAGTCTTCAGAGTTTTCAGGCAAGAAATCTTTTGAT 930
Db      329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348
QY      931 TTTATCGAAAACATTTCTTTAGAGGAAAACAAATCTTTGAGAAAACGAGTTTCAGAG 990
Db      349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368
QY      991 TATCAGCGTTTTCAGTTATGCGAGAAACACACAGATAACGCTCTTCACCTTGGATGCGAGAT 1050
Db      369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
QY      1051 TTT 1053
Db      389 Phe 389

RESULT 3
S06735
ribonucleoside-diphosphate reductase (EC 1.17.4.1) chain M2 - mouse
N/Alternate names: ribonucleotide reductase M2 subunit
C/Species: Mus musculus (house mouse)
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C/Accession: S06735; A24835
R/Thelander, M.; Thelander, L.
EMBO J. 8, 2475-2479, 1989
A/Title: Molecular cloning and expression of the functional gene encoding the M2 subunit
A/Reference number: S06735; MUID:90060004; PMID:2684652
A/Accession: S06735
A/Status: translation not shown
A/Molecule type: DNA
A/Residues: 1-390 <THE>
A/Cross-references: UNIPROT:P11157; EMBL:X15666; NID:g50719; PID:CAA33707.1; PID:g50720
R/Thelander, L.; Berg, P.
Mol. Cell. Biol. 6, 3433-3442, 1986
A/Title: Isolation and characterization of expressible cDNA clones encoding the M1 and M2 subunits of mouse ribonucleoside reductase
A/Reference number: A24835; MUID:87089677; PMID:3025593
A/Accession: A24835
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-390 <THE2>
A/Cross-references: GB:M14223; NID:g200767; PIDN:AAA40062.1; PID:g200768
A/Note: The authors translated the codon GGG for residue 315 as Glu
C/Genetics:
A/Map position: 12
A/Introns: 33/3; 58/3; 107/3; 146/3; 191/2; 223/1; 267/3; 302/3; 340/3
C/Superfamily: ribonucleoside reductase small subunit
C/Keywords: iron; oxidoreductase
F:139,170,173,233,267,270/Binding site: 2Fe-O cluster (Asp, Glu, His, Glu, His) #st
F:177/Active site: Tyr (stable tyrosyl radical) #status predicted

Alignment Scores:
Pred. No.: 6,58e-123 Length: 390
Score: 1443.00 Matches: 269
Percent Similarity: 90.62% Conservative: 40
Best Local Similarity: 78.89% Mismatches: 26
Query Match: 78.17% Indels: 6
DB: 1 Gaps: 1

US-10-698-228-2 (1-1053) x S06735 (1-390)

QY      49 AGATCATCTTCAGACACCAACGAAAGTGAATAAAGTCA-----AAAT 90
Db      50 ArgIlePheGlnAspSerAlaGluLeuGluSerLysAlaProThrAsnProSerValGlu 69
QY      91 GAAGAGCCACTCTTGAAGAAAGTTCTCCCGGTTTGTTCATCTTTCCAAATCCAGTACCTT 150
Db      70 AspGluProLeuLeuArgGluAsnProArgPheValValPheProIleGluTyrHis 89
QY      151 GATATTTCGAAATGTATTAACAGGCACAGGCTTCCTTCGACACGACGAGGCTCGAC 210

```

```

Db      90 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 109
QY      211 TTATCAAGAGATCCCTCTCACTCGAACCAAGCTTAAGCAGATGAGAGTACTTCTATCTCT 270
Db      110 LeuSerLysAspIleGlnHisTrpGluAlaLeuLysProAspGluArgHisPheIleSer 129
QY      271 CACATCTTAGCCCTTTTTCAGAGCCAGTCATGGAATTTGAAATGAAATTTTGGTGAGCGC 330
Db      130 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValLeuGluAsnLeuValGluArg 149
QY      331 TTTAGTCAGAGGTGTCAGGTTCCAGAGCTCGCTGTTCTTATGCTTTCAATTTCTCATC 390
Db      150 PheSerGlnGluValGlnValThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 169
QY      391 GAGAAATGTTCACTCAGAGATGTACAGTTTCTGTATAGACACTTACATCAGAGATCCCAAG 450
Db      170 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 189
QY      451 AAAAGGGAATTTTATTGAATTAATGCAATTAATGAAACCATGCTTATGTTAAGAAAAAGCAGAT 510
Db      190 GluArgGluTyrLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 209
QY      511 TGGGCTCTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 570
Db      210 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValValAlaPhe 229
QY      571 GCTGCTGTAGAAGAGTTTCTTCTCAGGATCTTTTCTGCTGCTATATTTCTGCTAAAGAG 630
Db      230 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 249
QY      631 AGAGGTCTTATGTCAGGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 690
Db      250 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 269
QY      691 CACTGTGACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 750
Db      270 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProAlaGluGlnArg 289
QY      751 GTCAGGAGATCATTTGTTGATGCTGCTCAAAATGAGCAGAGTTTAAACAGAGCCCTTG 810
Db      290 ValArgGluIleIleThrAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 309
QY      811 CCAGTTGGCTCATCTGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATG 870
Db      310 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 329
QY      871 GACAGATTACTTGTGGAATCTGGAATCTCAAGGTTTTCAGGAGAAAATCTTTTTCAT 930
Db      330 AspArgLeuMetLeuGluLeuGlyPheAsnLysIlePheArgValGluAsnProPheAsp 349
QY      931 TTTATCGAAAACATTTCTTTAGAGGAAAACAAATTTCTTTTGAAGAACCGAGTTTTCAGAG 990
Db      350 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 369
QY      991 TATCAGCGTTTTCAGTTATGCGAGAAACACACAGATAACGCTTTCACCTTCGATGCGAGAT 1050
Db      370 TyrGlnArgMetGlyValMetSerAsnSerThrGluAsnSerPheThrLeuAspAlaAsp 389
QY      1051 TTT 1053
Db      390 Phe 390

```

RESULT 4

```

S27153
ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain - golden hamster
N/Alternate names: ribonucleotide reductase small chain
C/Species: Mesocricetus auratus (golden hamster)
C/Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C/Accession: S27153
R/Chaudhuri, M.M.; Tonin, P.N.; Srinivasan, P.R.
Biochim. Biophys. Acta 1171, 117-121, 1992
A/Title: cDNA sequence of the small subunit of the hamster ribonucleotide reductase.
A/Reference number: S27153; MUID:93042000; PMID:1384717

```


A:Accession: S27153

A:Molecule type: mRNA

A:Residues: 1-386 <CHA>

A:Cross-references: UNIPROT:Q0561; EMBL:X68127; NID:g49649; PIDN:CAA48232.1; PID:g49650

C:Superfamily: ribonucleoside reductase small subunit

C:Keywords: oxidoreductase

Alignment Scores:
Pred. No.: 6.56e-121 Length: 386
Score: 1421.00 Matches: 266
Percent Similarity: 91.5% Conservative: 38
Best Local Similarity: 80.12% Mismatches: 20
Query Match: 76.98% Indels: 8
DB: 2 Gaps: 2

US-10-698-228-2 (1-1053) x S27153 (1-386)

```
QY 70 GAAAGTGAATAAAGTCAAAAT-----GAAAGCCACTCTTAAAGAAAGAGTTCT 117
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 59 GluSerLysValSerThrAsnProSerValGluAspGluProLeuLeuArgGluAsnPro 78
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 118 CGCGGTTTGTCTATCTTCCATCCAGTACCTGATATTGGAAATGATATAACAGGCA 177
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 79 ArgArgPheValValPheProIleGluTyrHisAspIleTrpLysLysAla 98
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 178 CAGGCTTCCTTCGGACACAGAGAGTTCGACTTATCAAGGATCTCCCTCACTGGAAC 237
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 99 GluAlaSerPheTrpThrAlaGluValAspLeuSerLysAspIleGlnHisTrpGlu 118
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 238 AAGCTTAAAGCAGATGAGAAGTACTTCATCTCTCACATCTTAGCCCTTTTTCGAGCCAGT 297
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 119 AlaLeuLysProAspGluArgHisPheIleSerHisValLeuAlaPhePheAlaAlaSer 138
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 298 GATGGAATTTGTAATCAAAATTTGGTGGAGCCCTTTAGTCAGAGGTGAGGTTCCAGAG 357
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 139 AspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnThrGlu 158
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 358 GCTCGCTGTTTCTATGGCTTCAAAATTCATCGAGAATGTCACACAGAGTGTACAGT 417
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 159 AlaArgCysPheTyrGlyPheGlnIleAlaMetGluAsnIleHisSerGluMetTyrSer 178
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 418 TTGCTGATAGACACTTACATCAGATGCCAAGAAAGGAAATTTTATTTATGCAATT 477
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 179 LeuLeuIleAspThrTyrIleLysAspSerLysGluArgGluTyrLeuPheAsnAlaIle 198
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 478 GAAACATCCCTATGTTAAGAAAAAGACAGATTTGGGCGCTTGGATGATAGCAGATAGA 537
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 199 GluThrMetProCysValLysLysAlaAspTrpAlaLeuArgTrpIleGlyAspLys 218
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 538 AAATCTACTTTTGGGAAAGAGTGGGCGCTTGTCTGTAGAGGAGTTTCTTCTCA 597
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 219 GluAlaThrTyrGlyGluArgValAlaPheAlaValGluGluIlePhePheSer 238
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 598 GGATCTTTTGTCTGCTATATCTGGCTTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTT 657
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 239 GlySerPheAlaSerIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPhe 258
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 658 TCAATGAATCATCAGCAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGCTGATGTC 717
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 259 SerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPhe 278
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 718 CAATACTTAGTAATAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTC 777
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 279 LysHisLeuValHisLysProSerGluGlnArgValGlnGluIleThrAsnAlaVal 298
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 778 AAAATGACGAGGAGTTTAAACAGAAAGCCTTGGCAGTTGGCTCATTTGGAATGTC 837
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 299 ArgIleGluGlnGluPheLeuThrGluAlaLeuProValLysLeuIleGlyMetAsnCys 318
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 838 ATTTTGTATGAACAGTACATGAGTTTGTAGCTGACAGATTAATCTGTGGAACCTTGGATTC 897
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 319 ThrLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuMetLeuGluLeuGlyPhe 338
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
QY 898 TCAAAGGTTTTTCAGGCAGAAAAATCCTTTTATGATTTTATGAAAAACATTTCTTTAAGAGA 957
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 339 AsnLysIlePheLysValGluAsnProPheAspPheMetGluAsnIleSerLeuGluGly 358
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 958 AAAACAAAATTTCTTTGAGAAACAGATTTTCAGATATCAGCGTTTTCAGATTATGCAGAA 1017
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 359 LysThrAsnPhePheGluLysArgValGlyGlnArgMetGlyValMetSer--- 377
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1018 ACCACAGATAACGCTTTCACCTTGGATGACAGATTTT 1053
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 378 -----AsnSerPheThrLeuAspAlaAspPhe 386
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
RESULT 5
S24585
ribonucleoside-diphosphate reductase (EC 1.17.4.1) chain M2 - Atlantic surf clam
N;Alternate names: ibonucleoside-diphosphate reductase small chain
C;Species: Spisula solidissima (Atlantic surf clam)
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: S24585
R;Hunt, T.; Standart, N.
submitted to the EMBL Data Library, November 1990
A;Reference number: S24584
A;Accession: S24585
A;Molecule type: mRNA
A;Residues: 1-384 <HUN>
A;Cross-references: UNIPROT:P07201; EMBL:X55125; NID:g10341; PIDN:CAA38919.1; PID:g10342
C;Superfamily: ribonucleoside reductase small subunit
C;Keywords: oxidoreductase
```

Alignment Scores:

Pred. No.:	9.93e-113	Length:	384
Score:	1331.00	Matches:	257
Percent Similarity:	85.26%	Conservative:	38
Best Local Similarity:	74.28%	Mismatches:	41
Query Match:	72.10%	Indels:	10
DB:	2	Gaps:	4

US-10-698-228-2 (1-1053) x S24585 (1-384)

```
QY 46 GAGAGATCATCTTCAGACACCAAC---GAAAGTGAATAAAG----- 84
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 39 GlnArgSerThrGlnLysThrProLeuLysGlnIleLysProValValLysLysSer 58
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 85 TCAATGAAGACGACCTCTTAAGAAAGAGTCTCGCGGTTTGTCACTTTTCCAAATCCAG 144
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 59 GlnGlnValGluProLeuLeuAlaAspAsnProArgPheValLeuProIleGln 78
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 145 TACCTCTGATATTTGGAAAAATGTATAACAGCGACAGGCTTCTCTCGACAGCAGAGAG 204
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 79 TyrHisAspIleTrpLysMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGlu 98
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 205 GTCGACTATCAAGGATCTCCCTCACTGGAAACAGCTTAAAGCAGATGAGAGTACTTTC 264
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 99 ValAspLeuSerLysAspMetAlaHisTrpGluSerLeuLysLysGluGluLysHisPhe 118
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 265 ATCTCTCACATCTTAGCTTTTTCAGCCAGGATGATGGAATGTAATGAAAATTTGGTG 324
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 119 IleSerHisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuVal 138
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 325 GAGCGCTTTAGTCAGGAGGTGCGAGGTTCCAGAGCTCGCTGTTTCTATGGCTTTCAAAT 384
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 139 GluArgPheSerLysGluValGlnValThrGluAlaArgCysPheTyrGlyPheGlnIle 158
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 385 CTCACTCAGAAATTTCACTCAGAGATGTACAGTTTGTGTATAGACACTTACATCAGAGAT 444
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 159 AlaMetGluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAsp 178
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 445 CCCAAGAAAGGAAATTTTATTAATGCAATGCAATGAAACCATGCCCTATGTTAAGAAAAA 504
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 179 ProGlnGluArgAspPheLeuPheAsnAlaIleGluThrMetProCysValLysGluLys 198
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 505 GCAGATTGGGCTTGGCATGATAGCAGATAGAAAAATCTACTTTTGGGAAAGAGTGTGTG 564
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

Db 199 AlaAepTAlaMetArgTrpIleAenAaspAepSerSerTyrAlaGluArgValVal 218
 QY 565 GCCTTTGCTGCTAGAACGAGTTTCTTCTCAGGATCTTTGCTGCTATATTTCTGGCTA 624
 Db 219 AlaPheAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeu 238
 QY 625 AAGAAGAGAGGCTTTATGCGCAGGACTCCTCTTTTCCAAATGAACTCATCAGCAGAGTGA 684
 Db 239 LysLysArgGlyMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAaspGlu 258
 QY 685 GCACTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 744
 Db 259 GlyLeuHisCysAaspPheAlaCysLeuMetPheSerHisLeuValAsnLysProSerGln 278
 QY 745 GAAAGGTCAGGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 804
 Db 279 GluArgIleHisGlnIleLeuAaspGluAlaValLysIleGluInValPheLeuThrGlu 298
 QY 805 GCCTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 864
 Db 299 AlaLeuProCysArgLeuIleGlyMetAsnCysAaspLeuMetArgGlnTyrIleGluPhe 318
 QY 865 GTAGCTGACAGTACTCTGGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 924
 Db 319 ValAlaAaspArgLeuLeuLeuLysCysAaspLysLeuTyrAsnLysGluAsnPro 338
 QY 925 TTTGATTTTATGGAACACTTCTTTAGAGGAAACAAATTTCTTTGAGAAACGAGTT 984
 Db 339 PheAaspPheMetGluHisIleSerLeuGluGlyLysThrAsnPhePheGluLysArgVal 358
 QY 985 TCAGAGTATCAGGCTTTTTCAGTTATGCGCA-----GAAACACACAGAT--AACGCTTC 1035
 Db 359 GlyGluTyrGlnLysMetGlyValMetSerGlyLysAsnThrGlyAaspSerHisAlaPhe 378
 QY 1036 ACCTTGGATGCAGATTTT 1053
 Db 379 ThrLeuAaspAlaAaspPhe 384

RESULT 6
 T30782
 ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain - vaccinia virus (strain
 N) Alternate names: ribonucleotide reductase, small subunit
 C:Species: vaccinia virus
 A:Variety: strain Ankara
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 R:Antoine, G.; Scheiflinger, F.; Falkner, F.G.; Dorner, F.
 submitted to the EMBL Data Library, March 1997
 A:Description: The complete genomic sequence of the Modified Vaccinia Ankara (MVA) strain
 A:Reference number: Z20877
 A:Accession: T30782
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-319 <ANT>
 A:Cross-references: UNIPROT:O57175; EMBL:U94848; PIDN:AAB96415.1
 C:Experimental source: strain Ankara
 C:Genetics:
 A:Note: MVA032L
 C:Function:
 A:Description: catalyzes the reduction of ribonucleoside diphosphate to deoxyribonucleoside
 C:Superfamily: ribonucleoside reductase small subunit
 C:Keywords: deoxyribonucleotide biosynthesis; oxidoreductase

Alignment Scores:
 Pred. No.: 4, 13e-108 Length: 319
 Score: 1280.00 Matches: 239
 Percent Similarity: 85.31% Conservatives: 34
 Best Local Similarity: 74.69% Mismatches: 45
 Query Match: 69.34% Indels: 2
 DB: 2 Gaps: 2

US-10-698-228-2 (1-1053) x T30782 (1-319)

QY 94 GAGCCACTCCTAAGAAAGAGTCTTCGCCGGTTTGTCTCATCTTTTCCAAATCCAGTACCCCTGAT 153
 Db 2 GluProIleLeuAlaProAenProAenArgPheValIlePheProIleGlnTyrHisAasp 21
 QY 154 ATTTGGAATAATGATATAACAGGACAGGCTTCTTCTGAGCAGCAGCAGAGAGGTCGACTTA 213
 Db 22 IleTrpAenMetTyrLysLysAlaGluAaSerPheTrpThrValGluGluValAaspIle 41
 QY 214 TCAAGAGTCTCCTCCTCAGTGAACAAAGCTTAAAGCAGATGAGAGTAGTCTTCATCTCTCAC 273
 Db 42 SerLysAaspIleAenAaspTrpAsnLysLeuThrProAaspGluLysTyrPheIleLysHis 61
 QY 274 ATCTTAGCTCTTTTTCAGCCAGTGTGAATGTAAATGAAATTTGGTGGAGCGCTTT 333
 Db 62 ValLeuAlaPhePheAlaSerAaspGlyIleValAsnGluAsnLeuAlaGluAargPhe 81
 QY 334 AGTCAGAGGTGCGGTTCAGAGGCTCGCTGCTTCTATGCTGCTTCAAAATTTCTCATCAG 393
 Db 82 CysThrGluValGlnIleThrGluAaArgCysPheTyrGlyPheGlnMetAlaIleGlu 101
 QY 394 AATGTTCACTCAGAGATGTACAGTTGCTGATAGACATTTACATCAGAGATCCCAAGAA 453
 Db 102 AsnIleHisSerGluMetTyrSerLeuLeuIleAaspThrTyrValLysAaspSerAsnGlu 121
 QY 454 AGGGAATTTTATTAATGCAATTGAAACCATGCCCTATGTTAAGAAAAAGCAGATGG 513
 Db 122 LysAsnTyrLeuPheAsnAlaIleGluThrMetProCysValLysLysLysAlaAaspTrp 141
 QY 514 GCCTTCGATGGATAGCAGATGAAATCTACTTTTGGGAAAGAGTGGTGGCTTGGCT 573
 Db 142 AlaGlnLysTrpIleHisAasp--SerAlaGlyTyrGlyGluArgLeuIleAlaPheAla 160
 QY 574 GCTGTAGAGAGGATTTTCTTCTCAGGATCTTTGCTGCTCTATATTTCTGGCTAAAGAGAGA 633
 Db 161 AlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysArg 180
 QY 634 GGTCTTATGCCAGGACTCACTTTTCCAATGAATCATCAGCAGAGATGAGGACTTCAC 693
 Db 181 GlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAaspGluGlyLeuHis 200
 QY 694 TGTGACTTTGCTGCTGATGTTTCCAATCTTAAATAGCCTTCAGAGAAAGGGTCT 753
 Db 201 CysAaspPheAlaCysLeuMetPheLysHisLeuHisProProSerGluThrVal 220
 QY 754 AGGAGATCATTTGATGCTGCTCAAAATTTAGCAGAGGATTTTAAACAGAGCTTGCCA 813
 Db 221 ArgSerIleIleThrAaspAlaValSerIleGluGlnGluPheLeuThrAlaAlaLeuPro 240
 QY 814 GTTGGCTCATTTGGAATGATTCATTTGATGAAACAGTACATTTGATTTGTAGCTGAC 873
 Db 241 ValLysLeuIleGlyMetAsnCysGluMetLysThrTyrIleGluPheValAlaAasp 260
 QY 874 AGATTACTTGTGGAATCTGGAATCTCAAGGTTTTCAGGCAGAGAAATCTTTTGATTTT 933
 Db 261 ArgLeuIleSerGluLeuGlyPheLysLysIleTyrAsnValThrAsnProPheAaspPhe 280
 QY 934 ATGGAACCATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACGAGTTTCAGAGPAT 993
 Db 281 MetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGluTyr 300
 QY 994 CAGCGTTTTCAGTTATGGCAGAAACCCACAGATAACGCTTTCACCTTGGATGACAGATTT 1053
 Db 301 GlnLysMetGlyValMetSerGln--GluAaspAsnHisPheSerLeuAaspValAaspPhe 319

RESULT 7
 B72154
 E4L protein - variola minor virus (strain Garcia-1966)
 C:Species: variola minor virus
 C:Date: 24-Nov-1999 #sequence_revision 24-Nov-1999 #text_change 09-Jul-2004
 C:Accession: B72154
 R:Shchelkunov, S.N.; Totmenin, A.V.; Gutorov, V.V.; Safronov, P.P.; Massung, R.F.; Lopar
 submitted to GenBank, March 1998
 A:Description: Analysis of the complete coding sequence of DNA of alastrim variola minor

A:Reference number: A72150
A:Accession: B72154
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-319 <SHC>
A:Cross-references: UNIPROT:Q89087; GB:Y16780; NID:G5830555; PIDN:CA54628.1; PID:G58305
A:Experimental source: strain Garcia-1966
C:Genetics:
A:Gene: E4L
C:Superfamily: ribonucleoside reductase small subunit

Alignment Scores:
Pred. No.: 4.13e-108 Length: 319
Score: 1280.00 Matches: 240
Percent Similarity: 85.31% Conservative: 33
Best Local Similarity: 75.00% Mismatches: 45
Query Match: 69.34% Indels: 2
DB: 2 Gaps: 2

US-10-698-228-2 (1-1053) x B72154 (1-319)

QY 94 GAGCCACTCTTAAGAAAGAGTTCTCGCGTTTGTCTATCTTCCATCCAGTACCCTGAT 153
Db 2 GluProIleLeuAlaLysAsnProAsnArgPheValIlePheProIleGlnTyrHisAsp 21

QY 154 ATTGGAAATGTATAACAGGCACAGGCTTCTCTGACAGCAGAGAGGTCCACTTA 213
Db 22 IleTyrAsnMetTyrLysLysAlaGluAlaSerPheTyrThrValGluGluValAspIle 41

QY 214 TCAAGGATCTCCCTCCTACCTGGAACAGCTTAAGCAGATGAGAAGTACTCTCTCAC 273
Db 42 SerLysAspIleAsnAspTyrAsnLysLeuThrProAspGluLysTyrPheIleLysHis 61

QY 274 ATCTTAGCTTTTTCGACCAAGTATGCAATGTAAATGAAATTTGGTGGCGCTTT 333
Db 62 ValLeuAlaPheAlaAlaSerAspGlyIleValAsnGluAsnLeuAlaGluArgPhe 81

QY 334 AGTCAGAGGTTCGAGGTTCACAGGCTCGCTGTTCTATGGCTTTCAAATTTCTCATCGAG 393
Db 82 CysIleGluValGluIleThrGluAlaArgCysPheTyrGlyPheGlnMetAlaIleGlu 101

QY 394 AATGTTCACTCAGAGATGACAGTTGCTGATAGACACTTACATCAGAGATCCCAAGAA 453
Db 102 AsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrValLysAspSerAsnGlu 121

QY 454 AGGGAATTTTATTAATGCAATTCGAACCATCCCTATGTTAAGAAAACAGAGTTGG 513
Db 122 LysAsnTyrLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAspTyr 141

QY 514 GCCTTGCGATGATGACAGATAGAAATCTACTTTGGGAAAGAGTGGTGGCTTTGCT 573
Db 142 AlaGlnLysTyrIleHisasp--SerAlaGlyTyrGlyGluArgLeuIleAlaPheAla 160

QY 574 GCTGTAGAGAGATTTTCTCTCAGGATCTTTGCTGCTATATCTTGGCTAAGAGAGA 633
Db 161 AlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTyrLeuLysLysArg 180

QY 634 GGTCTATGCCAGGACTCACTTTTCCAAATGAACTCATCAGCAGAGATGAAGACTTCAC 693
Db 181 GlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHis 200

QY 694 TGTGATTTGCTTGTCTGATGTTTCCAAATCTAGTAAATTAAGCCTTCAGAAAGGGTC 753
Db 201 CysAspPheAlaCysLeuMetPheLysHisLeuLeuTyrProProSerGluGluThrVal 220

QY 754 AGGAGATCATTTGTTGATGCTGCTCAAAATTTAGCAGGAGTTTTAAACAGAGCTTGCCA 813
Db 221 ArgSerIleIleThrAspAlaValSerIleGluGlnGluPheLeuThrValAlaLeuPro 240

QY 814 GTTGGCTCATTTGGAATGATTTGCAATTTTGTATGAAACAGTACATGAGTTTGTAGCTGAC 873
Db 241 ValLysLeuIleGlyMetAsnCysGluMetMetLysThrTyrIleGluPheValAlaAsp 260

QY 874 AGATTACTTGTGAACCTTGATTTCTCAAGGTTTTCAGCGAGAAATCCTTTTGATTTT 933
Db 261 ArgLeuIleSerGluLeuGlyPheLysLysIleTyrAsnValThrAsnProPheAspPhe 280

QY 934 ATGGAACATTTCTTTAGAGAAACAAATTTCTTTGAGAAACGAGTTTCAGAGTAT 993
Db 281 MetGluAsnIleSerLeuGluGlyLysThrAsnPheGluLysArgValGlyGluTyr 300

QY 994 CAGCGTTTTCAGTTATGCGAGAAACACAGATAAAGCTCTTCCACTTGGATGCAGATTTT 1053
Db 301 GlnLysMetGlyValMetSerGln---GluAspAsnHisPheSerLeuAspValAspPhe 319

RESULT 8
RDVZVV
ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain - vaccinia virus
N:Alternate names: P4L protein
C:Species: vaccinia virus
C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 09-Jul-2004
C:Accession: A29892; I42506; F36213
R:Slabaugh, M.; Roseman, N.; Davis, R.; Mathews, C.
J. Virol. 62, 519-527, 1988
A:Title: Vaccinia virus-encoded ribonucleotide reductase: sequence conservation of the
A:Reference number: A29892; MUID:88091062; PMID:2826813
A:Accession: A29892
A:Molecule type: DNA
A:Residues: 1-319 <SLA>
A:Cross-references: UNIPROT:P11158; GB:M19117; NID:G335808; PIDN:AAA88680.1; PID:G335809
A:Experimental source: strain WR
R:Goebel, S.O.; Johnson, G.P.; Perkus, M.E.; Davis, S.W.; Winslow, J.P.; Paoletti, E.
Virology 179, 517-563, 1990
A:Title: Appendix to "The complete DNA sequence of vaccinia virus".
A:Reference number: A42501
A:Accession: I42506
A:Molecule type: DNA
A:Residues: 1-212, 'Y', 214-319 <GOR>
A:Cross-references: GB:M35027; NID:G335317; PIDN:AAA48018.1; PID:G335366
A:Experimental source: strain Copenhagen
R:Goebel, S.J.; Johnson, G.P.; Perkus, M.E.; Davis, S.W.; Winslow, J.P.; Paoletti, E.
Virology 179, 247-266, 1990
A:Title: The complete DNA sequence of vaccinia virus.
A:Reference number: A42531; MUID:91021027; PMID:2219722
A:Contents: annotation; possible protein-coding frames
A>Note: neither amino acid nor nucleotide sequence is given
R:Roseman, N.A.; Slabaugh, M.B.
Virology 178, 410-418, 1990
A:Title: The vaccinia virus HindIII F fragment: nucleotide sequence of the left 6.2 kb.
A:Reference number: A36213; MUID:91020979; PMID:2219701
A:Accession: F36213
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-319 <ROS>
A:Cross-references: EMBL:M34368; NID:G335618; PIDN:AAA48244.1; PID:G335624
A:Experimental source: strain WR
C:Function:
A:Description: catalyzes the reduction of ribonucleoside diphosphate to deoxyribonucleos
C:Superfamily: ribonucleoside reductase small subunit
C:Keywords: DNA replication; iron; metalloprotein; oxidoreductase
F:70,104,163,197,200/Binding site: 2Fe-O cluster (Asp, Glu, His, Glu, His) #stat
F:108/Active site: Tyr (stable tyrosyl radical) #status predicted

Alignment Scores:
Pred. No.: 5.09e-108 Length: 319
Score: 1279.00 Matches: 239
Percent Similarity: 85.31% Conservative: 34
Best Local Similarity: 74.69% Mismatches: 45
Query Match: 69.28% Indels: 2
DB: 1 Gaps: 2

US-10-698-228-2 (1-1053) x RDVZVV (1-319)

QY 94 GAGCCACTCTTAAGAAAGAGTTCTCGCGTTTGTCTATCTTCCATCCAGTACCCTGAT 153
Db 2 GluProIleLeuAlaProAsnProAsnArgPheValIlePheProIleGlnTyrTyrAsp 21

Db 275 ArgLeuIleSerGluLeuGlyPheLysLysIleTyrAsnValThrAsnProPheAspPhe 294
QY 934 ATGGAAACATTTCTTAGAAGGAAACAAATTTCTTGGAGAAACGAGTTTCAGAGTAT 993
Db 295 MetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGluTyr 314
QY 994 CAGCGTTTGGCAGTTATGGCAGAAACCCAGATAACGTCCTTCCACCTTGGATCGACATTTT 1053
Db 315 GlnLysMetGlyValMetSerGln---GluAspAsnHisPheSerLeuAspValAspPhe 333
RESULT 10
H36839
CBU protein - variola virus (strain India-1967)
C:Species: variola virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C:Accession: H36839
R:Blinov, V.M.
submitted to GenBank, November 1992
A:Reference number: A36859
A:Accession: H36839
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-333 <BLI>
A:Cross-references: UNIPROT:P33799; GB:X69198; NID:g456758; PIDN:CAA48969.1; PID:g297209
C:Superfamily: ribonucleoside reductase small subunit
Alignment Scores:
Pred. No.: 7,81e-107 Length: 333
Score: 1266.00 Matches: 237
Percent Similarity: 85.00% Conservative: 35
Best Local Similarity: 74.06% Mismatches: 46
Query Match: 68.58% Indels: 2
DB: 2 Gaps: 2
US-10-698-228-2 (1-1053) x H36839 (1-333)
QY 94 GAGCCACTCTTAAGAAAGAGTTCTCGCCGGTTGTTCATCTTTCCAATCCAGTACCTGTAT 153
Db 16 GluProIleLeuAlaLysAsnProAsnArgPheValIlePheProIleGlnTyrHisAsp 35
QY 154 ATTGTGAAATATATAACAGCACAGGCTTCTTCTGGACAGCAGAGAGTCGATTA 213
Db 36 IleTyrAsnMetTyrLysLysAlaGluAlaSerPheTyrThrValGluGluValAspIle 55
QY 214 TCAAGAGTCTCCCTCACTCGACAGCTTAAGACAGATGAGAGTACTTCTCTCTAC 273
Db 56 SerLysAspIleAsnAspTrpAsnLysLeuThrProAspGluLysTyrPheIleLysHis 75
QY 274 ATCTTAGCCTTTTTCAGCCAGTGTGAATTTGAAATTTGGTGAGCGCTTT 333
Db 76 ValLeuAlaPheAlaAlaSerAspGlyIleValAsnGluAsnLeuAlaGluArgPhe 95
QY 334 AGTCAGGAGGTCAGGTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAAATTCATCGAG 393
Db 96 CysIleGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnMetAlaIleGlu 115
QY 394 ATGTTCTACTCAGAGTGTACAGTTGTGCTGATAGACACTTACATCAGATCCCAAGAA 453
Db 116 AsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrValLysAspSerAsnGlu 135
QY 454 AGGGAATTTTATTAAATCAATGAAACCATGCCCCTATGTTAAAGAAACCAAGCATTCG 513
Db 136 LysAsnTyrLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAspTrp 155
QY 514 GCCTTGGCATGATGACAGATAAATCTACTTTTGGGAAAGAGTGGTGCTTTGTCT 573
Db 156 AlaGlnLysTrpIleHisAsp---SerAlaGlyTyrGlyGluArgLeuIleAlaPheAla 174
QY 574 GCTGTAGAGAGGATTTTCTTCAGAGATCTTTGCTGCTATATTCGGCTTAAGAGAGA 633
Db 175 AlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLysArg 194
QY 634 GGTCTTATGCCAGGACTCACTTTTCCATGAACTCATCAGCAGATGAAGGACTTCAC 693

Db 195 GlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHis 214
QY 694 TGTGACTTTCCTTCCTCGATGTTCCAACTACTAGTAATAAGCCTTCAGAGAAGGGTC 753
Db 215 CysAspPheAlaCysLeuMetPheLysHisLeuTyrProProSerGluGluThrVal 234
QY 754 AGGGAGATCATTTGTTGATGCTGTCAAAATTTGAGCAGGAGTTTAAACAGAACCTTGGCA 813
Db 235 ArgSerIleIleThrAspAlaValSerIleGluGlnGluPheLeuThrValAlaLeuPro 254
QY 814 GTTGGCCTCATTTGGAATGAATTCATTTTGTGATGAACAGACATCATGTAGTTGTAGCTGAC 873
Db 255 ValLysLeuIleGlyMetAsnCysGluMetMetLysThrTyrMetGluPheValAlaAsp 274
QY 874 AGATTACTTGTGGAACCTTGGATTTCTCAAGGTTTTCAGGCAGAAAATCCTTTTGATT 933
Db 275 ArgLeuIleSerGluLeuGlyPheLysArgIleTyrAsnValThrAsnProSerAspPhe 294
QY 934 ATGGAAACATTTCTTTAGAGGAAACAAATTTCTTTCAGAAACGAGTTTCAGAGTAT 993
Db 295 MetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGluTyr 314
QY 994 CAGCGTTTTCAGTTATGGCAGAAACCCAGATAACGTCCTTCCACCTTGGATCGACATTT 1053
Db 315 GlnLysMetGlyValMetSerGln---GluAspAsnHisPheSerLeuAspValAspPhe 333
RESULT 11
T30470
ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain - Lymantria dispar nuclear
N:Alternate names: ribonucleotide reductase small subunit homolog
C:Species: Lymantria dispar nuclear polyhedrosis virus, LdMNPV
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T30470
R:Kuzio, J.; Pearson, M.N.; Harwood, S.H.; Funk, C.J.; Evans, J.T.; Slavicek, J.M.; Rohrer
Virology 253, 17-34, 1999
A:Title: Sequence and analysis of the genome of a baculovirus pathogenic for Lymantria d
A:Reference number: Z20836; MUID:99124785; PMID:9887315
A:Accession: T30470
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-348 <KUZ>
A:Cross-references: UNIPROT:Q9YMK7; EMBL:AF081810; PIDN:AAC70306.1
C:Function:
A:Description: catalyzes the reduction of ribonucleoside diphosphate to deoxyribonucleos
C:Superfamily: ribonucleoside reductase small subunit
C:Keywords: oxidoreductase; pyrimidine deoxynucleotide metabolism
Alignment Scores:
Pred. No.: 3,78e-106 Length: 348
Score: 1258.50 Matches: 246
Percent Similarity: 80.63% Conservative: 37
Best Local Similarity: 70.09% Mismatches: 59
Query Match: 68.17% Indels: 9
DB: 2 Gaps: 3
US-10-698-228-2 (1-1053) x T30470 (1-348)
QY 1 ATGGCGCACCCGGAAGCGCGGAGCGCGCGCTGGATCAGGATGAGATCATCTTCA 60
Db 7 MetProAlaProGluArgPro-----PheAspProArgAlaProAla 20
QY 61 GACACCAACGAAGTAGTAATAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
Db 21 Pro-----GluArgProPheAspProArgAlaGluProLeuLeuArgGluAsnProArg 38
QY 121 CGTTTGTTCATCTTCCAACTCCAGTACCTCGATATTTGGAAATGTATAACAGCACAG 180
Db 39 ArgPheValIlePheProIleGlnTyrProAspMetTrpArgMetTyrLysLysAlaGlu 58
QY 181 GCTTCTCTTGGACAGCAGAGAGGTGACTTATCAAGGATCTCCCTCACTGGAACAA 240
Db 59 AlaSerPheTrpThrValGluValAspLeuSerLysAspThrSerAspTrpGluArg 78

C;Superfamily: ribonucleoside reductase small subunit

C;Keywords: iron; oxidoreductase; pyrimidine deoxynucleotide metabolism
F;145,176,179,239,273,276/Binding site: 2Fe-O cluster (Asp, Glu, His, Glu, His) #54
F;183/Active site: Tyr (stable tyrosyl radical) #status predicted

Alignment Scores:
Pred. No.: 4,78e-98 Length: 399
Score: 1169.50 Matches: 221
Percent Similarity: 81.07% Conservative: 53
Best Local Similarity: 65.38% Mismatches: 58
Query Match: 63.35% Indels: 6
DB: 1 Gaps: 3

US-10-698-228-2 (1-1053) x A26916 (1-399)

```

QY 54 ATCTTCAGACACCAACGAAGTGAATAAGTCAATGAAGCCACTCTTAGAAGAG 113
Db 64 ValHisArgHisLeuLysLeuLys-GluMetGlu---LysGluGluProLeuLeuAsnGluAs 82
QY 114 TTCTCCCGGTTTGTTCATCTTCCATCCAGTACCTCGATATTGGAAATGTATAACA 173
Db 82 physGluArgThrValLeuPheProIleLysTyrHisGluIleTrpGlnAlaTyrLysAr 102
QY 174 GGCACAGGCTTCTTCGACGACGACGAGGTCGACTTATCAAGAGATCTCCCTCACTG 233
Db 102 gAlaGluAlaSerPheTrpThrAlaGluGluLeuAspLeuSerLysAspIleHisAspTr 122
QY 234 G---AACAGCTTAAGACGATGAGAGTACTTCTCATCTCTCATCTTACGCTTTTTCG 290
Db 122 pAsnAsnArgMetAsnGluAsnGluArgPhePheIleSerArgValLeuAlaPheAl 142
QY 291 AGCCAGTGTGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGT 350
Db 142 aAlaSerAspGlyIleValAsnGluAsnLeuValGluAsnPheSerThrGluValGlnIl 162
QY 351 TCAGAGGCTCGCTGTTTCTATGGCTTTTCAAAATTTCTCATCGAAATGTTTCACCTCAGAGAT 410
Db 162 eProGluAlaLysSerPheTyrGlyPheGlnIleMetIleGluAsnIleHisSerGluTh 182
QY 411 GTACAGTTGCTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAA 470
Db 182 rTyrSerLeuLeuIleAspThrTyrIleLysAspProLysGluSerGluPheLeuPheAs 202
QY 471 TCAATTGAACACCTGCTATGTTAAGAAAGCAGAGATGGCGCTTCGATGATAGC 530
Db 202 nAlaIleHisThrIleProGluIleGlyGluLysAlaGluTrpAlaLeuArgTrpIleGl 222
QY 531 AGATGAATAATCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAAGAGTTT 590
Db 222 nAspAlaAspAlaLeuPheGlyGluArgLeuValAlaPheAlaSerIleGluGlyValPh 242
QY 591 CTTCTCAGATCTTTGCTGCTATATCTGGCTAAAGAGAGAGGTCCTTATGCCAGACT 650
Db 242 ePheSerGlySerPheAlaSerIlePheTrpLeuLysLysArgGlyMetMetProGlyLe 262
QY 651 CACTTTTCCATGACTCATCAGCAGATGAGGACTTCATCTGCTGCTGCTGCTGCT 710
Db 262 uThrPheSerAsnGluLeuIleCysArgAspGluGlyLeuHisThrAspPheAlaCysLe 282
QY 711 GATGTTTCCAAATCTTAGTAAATAAGCTTTCAGAAAGAGGGTCAGGAGATCATTTGTA 770
Db 282 uLeuPheAlaHisLeuLysAsnLysProAspProAlaIleValGluLysIleValThrGl 302
QY 771 TGCTGTCAAAATGACGAGGAGTTTAAACAGAAGCTTTGCCAGTTGGCTCATTTGGAAT 830
Db 302 uAlaValGluIleGluGlnArgTyrPheLeuAspAlaLeuProValAlaLeuLeuGlyMe 322
QY 831 GAATTGCATTTTGTGAAACAGTACATGAGTTTGTAGCTGACAGATTACTTGTGGAAT 890
Db 322 tAsnAlaAspLeuMetAsnGlnTyrValGluPheValAlaAspArgLeuLeuValAlaPh 342
QY 891 TGGATTCTCAAGGTTTTCAGCGAGAAATCTCTTTGATTTTATGAAACATTTCTTT 950

```

```

Db 342 eGlyAsnLysLysTyrTyrLysValGluAsnProPheAspPheMetGluAsnIleSerLe 362
QY 951 AGAAGGAAAAACAATTTCTTTGAAAAACGAGTTTCAGAGTATCAGCGCTTTTGCAAGTTAT 1010
Db 362 uAlaGlyLysThrAsnPhePheGluLysArgValSerAspTyrGlnLysAlaGlyValMe 382
QY 1011 GGCAGAAACACAGAT-----AACGCTTTCACCTTTCGATGCGAGATTTT 1053
Db 382 tSerLysSerThrLysGlnGluAlaGlyAlaPheThrPheAsnGluAspPhe 399
RESULT 15
T39992
ribonucleoside-diphosphate reductase small chain - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T39992
R;lyne, M.; Rajandream, M.A.; Barrrell, B.G.; Whithead, S.; Chillingworth, T.; Churcher,
submitted to the EMBL Data Library, July 1998
A;Reference number: Z21897
A;Accession: T39992
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-391 <LYN>
A;Cross-references: UNIPROT:P36603; EMBL:AL031158; PIDN:CAA20100.1; GSPDB:GN00067; SPDB:
A;Experimental source: strain 972h-; cosmid c25D12
C;Genetics:
A;Gene: SPDB:SPBC25D12.04
A;Map position: 2
C;Superfamily: ribonucleoside reductase small subunit
Alignment Scores:
Pred. No.: 2,54e-97 Length: 391
Score: 1161.50 Matches: 227
Percent Similarity: 78.80% Conservative: 48
Best Local Similarity: 65.04% Mismatches: 61
Query Match: 62.92% Indels: 13
DB: 2 Gaps: 4
US-10-698-228-2 (1-1053) x T39992 (1-391)
QY 37 GATCAGGATGAGGATCATCTTTCAGAC-----ACCAACGAAAGTGAATAAAG 84
Db 46 GluGluLysGlnGluGlyAspTyrTyrLeuGlyLysLysGluAspGluLeu--- 64
QY 85 TCAATGAAGACGACCTCTTAAGAAAGAGTCTCGCGGTTTGTCACTTTCCATCCAG 144
Db 65 -----AspGluValValLeuArgProAsnProHisArgPheValLeuPheProIleLys 82
QY 145 TACCTCTGATATTGGAAATGTATAACAGCAGCAGGCTTCTTCTGACAGCAGAGAG 204
Db 83 TyrHisGluIleTrpGlnPheTyrLysLysAlaGluAlaSerPheTrpThrAlaGlu 102
QY 205 GTCGACTTATCAAGGATCTCCCTCACTGG---AACAAAGCTTAAAGCAGATGAGAAGTAC 261
Db 103 IleAspLeuSerLysAspLeuValAspTrpAspAsnLysLeuAsnAlaAspGluArgTyr 122
QY 262 TTCATCTCTCATCTTAGCTTTTTCGACGACGATGATGGAATGTAATAAATTTG 321
Db 123 PheIleSerThrValLeuAlaTyrPheAlaAlaSerAspGlyIleValAsnGluAsnLeu 142
QY 322 GTGAGCGCTTTAGTCAGAGGTGCGAGTTCAGAGGCTCGCTGCTTTCTATGGCTTCA 381
Db 143 LeuGluArgPheSerSerGluValGlnIleProGluAlaArgCysValTyrGlyPheGln 162
QY 382 ATTCTCATCGAGATGTTTCACTCAGAGATGTCAGAGTTGCTGTAGTACACACTTATCATCAGA 441
Db 163 IleMetIleGluAsnIleHisSerGluThrTyrSerLeuLeuLeuAspThrTyrIleArg 182
QY 442 GATCCCAAGAAAGGGAATTTTATTATGCAATTGAAACCATGCGCCTATGTTAAGAAA 501
Db 183 GluProLysGluLysGlnArgHisPheAspAlaIleLeuThrMetGlySerIleLysAla 202
QY 502 AAAGCAGATTGGCGCTTGGCATGGATAGACAGATAAGAAATCTACTTTTGGGGAAGAGTG 561

```


											:::	
Db	203	LysAlaLysTrpAlaLeuArgTfPrlIleAsnAspGluAspSerThrTyRAlaIleArgLeu	222									
QY	562	GTGGCTTCCTCGTGAGAAAGAGTTTCTCTCAGCATCTTTGTGCTATATTCTGG	621									
Db	223	ValAlaPheAlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTfp	242									
QY	622	CTAAGAAGAGAGGTCTTATGCCAGGACTCACATTTTTCCCAATGAATCATCAGCAGAGAT	681									
Db	243	LeuLysLysArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleCysArgAsp	262									
QY	682	GAAGGACTTCACGTGTGACTTTGTGCTGCATGTTCCAATTACTTAGTAATAAGCCCTCA	741									
Db	263	GluGlyLeuHisThrAspPheAlaCysLeuMetPheSerHisLeuLysHisArgProGly	282									
QY	742	GAAGAAAGGTCAGGAGAGATCATTTGTGATCGTGTCAAATTTGAGCAGGAGTMTTAAACA	801			:	:	:	:	:	:	:
Db	283	ArgLysValValGluAlaIleIleValGluAlaValAspIleGluLysGluTyPheThr	302			:	:	:	:	:	:	:
QY	802	GAAGCCTTGCAGTGGCCCTCATTTGGAATGCAATTCATTTTGATGAACAGTACATTGAG	861			:	:	:	:	:	:	:
Db	303	AspAlaLeuProValSerLeuLeuGlyMetAsnLysAspLeuMetCysGlnTyIleGlu	322			:	:	:	:	:	:	:
QY	862	TTTGTAAGTCACAGATTACTTTGTGGAATCTTGGAATCTCAAAGTMTTTCAGGCAGAAAAT	921			:	:	:	:	:	:	:
Db	323	PheValAlaAspArgLeuLeuValAlaLeuGlyAsnAspLysTyTyTyAsnValThrAsn	342			:	:	:	:	:	:	:
QY	922	CCTTTTCATTTTATGAAAAACATTTCTTTAGAAGGAAAAACAAATTTCTTTGAGAACGA	981			:	:	:	:	:	:	:
Db	343	ProPheAspPheMetGluAsnIleSerLeuAlaGlyLysThrAsnPhePheGluLysLys	362			:	:	:	:	:	:	:
QY	982	GT TTCAGAGTATCAGCGTTTTTG CAGTTATG-----GCGAAAAACCACAGAT	1056			:	:	:	:	:	:	:
Db	363	ValSerAspTyrglnIleAlaGlyValMetSerGlyThrLysArgAlaGluLysAspAsp	382			:	:	:	:	:	:	:
QY	1027	AACGTCCTTCACCTTGGATGCAGATTTT	1053			:	:	:	:	:	:	:
Db	383	HISThrPheThrIleAspGluAspPhe	391			:	:	:	:	:	:	:

Search completed: October 30, 2005, 06:35:01
Job time : 25.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 30, 2005, 03:28:20 ; Search time 76 Seconds
(without alignments)
14189.984 Million cell updates/sec

Title: US-10-698-228-2
Perfect score: 1846
Sequence: 1 atggcgaccggaaggcc.....tcacctggatgcagatttt 1053

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 3224756

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10698228/runat_26102005_100608_3981/app_query.fasta.1.2446
-DB=UniProt_03 -QMT=fastan -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10698228@cgn.1.149@runat_26102005_100608_3981 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt 03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	1821	98.6	351	2	Q7LG56	Q7LG56	homo sapien
2	1817	98.4	351	2	Q9NUW3	Q9nuw3	homo sapien
3	1725	93.4	351	2	Q6PEE3	Q6pee3	mus musculus
4	1534	83.1	297	2	Q6AI41	Q6ai41	homo sapien
5	1514	82.0	299	2	Q75PY9	Q75py9	homo sapien
6	1474	79.8	285	2	Q8VEB3	Q8veb3	homo sapien
7	1454	78.8	389	1	R1R2_HUMAN	P31350	homo sapien
8	1443	78.0	390	1	R1R2_MOUSE	P11157	mus musculus
9	1439	78.0	386	2	Q68EP1	Q68ep1	xenopus tro
10	1431	77.5	386	2	Q80104	Q801q4	xenopus lae
11	1427	77.3	386	2	Q8AVX2	Q8avx2	xenopus lae
12	1425.5	77.2	386	1	R1R2_BRARE	P79733	brachydanio
13	1421	77.0	386	1	R1R2_MESAU	Q60561	mesocricetu
14	1420.5	77.0	386	2	Q6D144	Q6d144	brachydanio
15	1399.5	75.8	378	2	Q6P876	Q6p876	xenopus tro
16	1385	75.0	406	2	Q6IP47	Q6ip47	xenopus lae

17	1382.5	74.9	388	2	Q7PF28	Q7pf28	anopheles g
18	1382.5	74.9	426	2	Q7QIF4	Q7qif4	anopheles g
19	1370	74.2	349	2	Q7ZVW0	Q7zyw0	brachydanio
20	1358	73.6	397	2	Q95VP8	Q95vp8	aedes aegypt
21	1349.5	73.1	399	2	Q27124	Q27124	urechis cau
22	1348	73.0	393	1	R1R2_DROME	P48592	drosophila
23	1335	72.3	398	2	Q9XYN8	Q9xyN8	aedes albop
24	1331	72.1	384	1	R1R2_SPI50	P07201	spisula sol
25	1284	69.6	403	2	Q6CFU6	Q6cfu6	yarrowia li
26	1280	69.3	319	2	O57175	O57175	vaccinia vi
27	1280	69.3	319	2	Q76Q46	Q76q46	variola min
28	1280	69.3	319	2	Q89087	Q89087	variola vir
29	1279	69.3	319	1	R1R2_VACCV	P11158	vaccinia vi
30	1279	69.3	319	2	Q76ZX1	Q76zx1	vaccinia vi
31	1277	69.2	319	2	Q6RZQ9	Q6rZq9	rabbitpox v
32	1276	69.1	319	1	R1R2_VACCC	P20493	vaccinia vi
33	1276	69.1	319	2	Q8JLH6	Q8jLh6	ectromelia
34	1276	69.1	319	2	Q89559	Q89559	variola vir
35	1274	69.0	333	2	Q8QN14	Q8qn14	cowpox viru
36	1273	69.0	319	1	R1R2_VACCP	P29883	vaccinia vi
37	1271	68.9	319	2	Q8V544	Q8v544	monkeypox v
38	1270	68.8	319	2	P87632	P87632	cowpox viru
39	1268	68.7	319	2	Q8V2X8	Q8v2x8	camelpox vi
40	1268	68.7	319	2	Q775Z5	Q775z5	camelpox vi
41	1267	68.6	319	2	Q9JFE9	Q9jfe9	vaccinia vi
42	1266	68.6	319	1	R1R2_VARV	P33799	variola vir
43	1264.5	68.5	410	1	R1R2_NEUCR	Q9c167	neutrospora
44	1260	68.3	333	2	Q9PXK4	Q9pxk4	variola vir
45	1258.5	68.2	348	2	Q9YMK7	Q9ymk7	lymantria d

ALIGNMENTS

RESULT 1

Q7LG56 PRELIMINARY; PRT; 351 AA.
ID Q7LG56 Q9NPDS; Q9NTD8;
AC Q7LG56 Q9NPDS; Q9NTD8;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Ribonucleotide reductase (Hypothetical protein DKFZp761E1312) (P53-
inducible ribonucleotide reductase small subunit 2).
GN Name=p53R2; Synonyms=DKFZp761E1312;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20179179; PubMed=10716435; DOI=10.1038/35003506;
RA Tanaka H., Arakawa H., Yamaguchi T., Shiraishi K., Fukuda S.,
RA Matsui K., Takei Y., Nakamura Y.,
RT "A ribonucleotide reductase gene involved in a p53-dependent cell-
cycle checkpoint for DNA damage.";
RL Nature 404:42-49(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Amygdala;
RG The German cDNA Consortium;
RA Ansorge W., Krieger S., Regiert T., Rittmuller C., Schwager B.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Ugai H., Yokoyama K.K.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB036532; BAA92493.1; JOINED.
DR EMBL; AB036524; BAA92493.1; JOINED.
DR EMBL; AB036525; BAA92493.1; JOINED.
DR EMBL; AB036526; BAA92493.1; JOINED.
DR EMBL; AB036527; BAA92493.1; JOINED.
DR EMBL; AB036528; BAA92493.1; JOINED.
DR EMBL; AB036529; BAA92493.1; JOINED.

DR GO; GO:0004748; F:ribonucleoside-diphosphate reductase activity; IEA.
DR GO; GO:0009186; P:deoxyribonucleoside diphosphate metabolism; IEA.
DR InterPro; IPR009078; Ferritin/RR like.
DR InterPro; IPR000358; Ribonucleotide reductase.
DR Pfam; PF00268; Ribonucleotide reductase.
DR PROSITE; PS00368; RIBORED_SMALL; 1.
SQ SEQUENCE 351 AA; 40704 MW; 6D009B2D59E9A323 CRC64;

Alignment Scores:
Pred. No.: 9.34e-151 Length: 351
Score: 1817.00 Matches: 350
Percent Similarity: 100.00% Conservativeness: 1
Best Local Similarity: 99.72% Mismatches: 0
Query Match: 98.43% Indels: 0
DB: 2 Gaps: 0

US-10-698-228-2 (1-1053) x Q9NUW3 (1-351)

Qy 1 ATGGGCGACCGGAAAGCCGCGGCTGGATCAGGATGAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20

Qy 61 GACACCAACGAAAGTGAATTAAGTCAATGAAGCCACCTCTTAAGAAAGATTCTCGC 120
Db 21 AspThrAsnGluSerGluLeuLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40

Qy 121 CGGTTTCGTCATCTTCCAAATCCAGTACCTGATATTTGGAAATCTATAACACGACGAC 180
Db 41 ArgPheValIlePheProIleGlnTyProAspIleTrpLysMetTyLysGlnAlaGln 60

Qy 181 GCTTCTCTTCGACGACGAGAGGCTCGACTTATCAAGAGGATCTCCCTCACTGGAACAAG 240
Db 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80

Qy 241 CTTAAACGACATGAAAGTACTTCTCATCTCTCAGTCTTTTTCGACCCAGTGCAT 300
Db 81 LeuLysAlaAspGluLysTyPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100

Qy 301 GGAATTCGAATGAATTTGGTGGAGCGCTTATGTCAGGAGTCCAGGTTCCAGAGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120

Qy 361 CGCTGTTCTTATGGCTTCAAAATCTCATCGAGATGTTCTACTCAGAGATGTACAGTTTG 420
Db 121 ArgCysPheTyGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTySerLeu 140

Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGGAATTTTATTTAATGCAATTCAA 480
Db 141 LeuIleAspThrTyIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160

Qy 481 ACCATGCCCTATGTTAAGAAAGAGATGGCGCTTGCATGATGATGATGATGATGATGAT 540
Db 161 ThrMetProTyValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180

Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCTTCTGCTGCTGTAGAGAGTTCCTCTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200

Qy 601 TCTTTTCTGCTATATTCTGGCTAAAGAGAGGTCTTATGCCAGGACTCACATTTTTC 660
Db 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220

Qy 661 AATGAACATCAGCAGATGAAGGACTTCACTGTGACTTTGCTGCTGCTGCTGCTGCTCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240

Qy 721 TACTTAGTAATAGCTTCAAGAAAGAGGTTCAGGAGATCATTTGTTGATGCTGTCAAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260

Qy 781 ATTGAGCAGGAGTTTAAACAGAGCCCTTGCAGTGGCCCTCATTTGGAATGATTCATT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyValAsnCysIle 280

Qy 841 TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGAACTTGGATTCTCA 900
Db 281 LeuMetLysGlnTyIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300

Qy 901 AAGTTTTTCAGGCGAGAAATCCTTTGATTGTTTATGGAACAACATTTCTTTAGAAGGAAA 960
Db 301 LysValPheGlnAlaGluAsnPropheAspPheMetGluAsnIleSerLeuGluGlyLys 320

Qy 961 ACAATTTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGCTTATGCGAGAAC 1020
Db 321 ThrAsnPhePheGluLysArgValSerGluTyGlnArgPheAlaValMetAlaGluThr 340

Qy 1021 ACAGATAACGCTTTCACCTTGGATGCGAGATT 1053
Db 341 ThrAsnValPheThrLeuAspAlaAspPhe 351

RESULT 3
Q6PEE3 PRELIMINARY; PRT; 351 AA.

AC Q6PEE3; TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Ribonucleotide reductase M2 B (TP53 inducible).
GN Name=Rm2b;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Strausberg R.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC058103; AAH58103.1; -
DR GO; GO:0004748; F:ribonucleoside-diphosphate reductase activity; IEA.
DR GO; GO:0009263; P:deoxyribonucleotide biosynthesis; IEA.
DR InterPro; IPR009078; Ferritin/RR like.
DR InterPro; IPR000358; Ribonucleotide reductase.
DR Pfam; PF00268; Ribonucleotide reductase.
DR PROSITE; PS00368; RIBORED_SMALL; 1.
SQ SEQUENCE 351 AA; 40803 MW; 4E1259233C9C8A9 CRC64;

Alignment Scores:
Pred. No.: 1.13e-142 Length: 351
Score: 1725.00 Matches: 328
Percent Similarity: 97.44% Conservativeness: 14
Best Local Similarity: 93.45% Mismatches: 9

Query Match: 93.45% Indels: 0
DB: 2 Gaps: 0
US-10-698-228-2 (1-1053) x Q6PBE3 (1-351)

QY 1 ATGGGACCCGGAAGCCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
DB 1 MetGlyAspProGluArgProGluAlaAlaArgProGluLysGlyGluLeuCysSer 20
61 GACACCAACGAAGTGAATAAGCTCAATGAAGAGCCACTCCCTAAAGAAAGAGTTCTCGC 120
DB 21 GluThrGluAsnValValArgSerAsnGluGluProLeuLeuArgLysSerSerArg 40
QY 121 CGGTTTGTCTATCTTCCAAATCCAGTACCCTGATATTTGGAAAATGTATAAACAGGCACAG 180
DB 41 ArgPheValIlePheProIleGlnTyrProAspIleIleTrpArgMetTyrLysGlnAlaGln 60
QY 181 GCTTCCTCTCGGACAGCAGAGAGGTGCAGTATCAAGGATCTCCCTCACTGGAACAG 240
DB 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAGACGATGAGAAAGTACTCTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTCAT 300
DB 81 LeuLysSerAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY 301 GGAATTGTAAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTTCCAGAGCCT 360
DB 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATGGCTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
DB 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTATTAATGCAATTGAA 480
DB 141 LeuIleAspThrTyrIleArgAspProLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTTATGTTAAGAAAAAGCAGATTTGGCGCTTGGCATGGATAGCAGATAGAAAA 540
DB 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTGGGAAAGAGTGGTGGCTTGTCTGCTGTAGAAAGAGTTTCTTCTCAGGA 600
DB 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyIlePhePheSerGly 200
QY 601 TCTTTTGTCTATATTCGCTAAAGAGAGTCTTATGCCAGGACTCACATTTTTC 660
DB 201 SerPheAlaAlaIlePheTrpLeuLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAATCATCAGCAGATGAGGACTTCACCTGTGACTTTGCTGCTCATGATGTTCCAA 720
DB 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAATAAGCCTTCAGAGAAGAGGTGAGGAGATCATGTTGTGATGCTGTCAAA 780
DB 241 TyrLeuValAsnLysProSerGluAspArgValArgGluIleIleAlaAspAlaValGln 260
QY 781 ATTGACGAGGTTTAAACAGAGCCTTCCAGTTCGCTCATTTGGAATGAATTCGATT 840
DB 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysVal 280
QY 841 TTGATGAAACAGATACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATCTCA 900
DB 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuGlyGluLeuGlyPheSer 300
QY 901 AAGGTTTTTCAGCAGAAAATCCTTTTGTATTTATGAAAAACATTTCTTTAGAGGAAAA 960
DB 301 LysIlePheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAAAATTTCTTCAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAC 1020
DB 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340

QY 1021 ACAGATAACGTTCTTCCACCTTGGATCAGATTTT 1053
DB 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 4
Q6AI41 PRELIMINARY; PRT; 297 AA.
AC Q6AI41;
DT 25-OCT-2004 (TRENBLrel. 28, Created)
DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)
GN Hypothetical protein DKFP686M05248.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RP The German cDNA Consortium;
RG The German cDNA Consortium;
RA Ottenwaelder B., Obermaier B., Deutschenbaur S., Schaipp A.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR627376; CAH10473.1; -;
DR GO; GO:0004748; F:ribonucleoside-diphosphate reductase activity; IEA.
DR GO; GO:0009186; P:deoxyribonucleoside diphosphate metabolism; IEA.
DR InterPro; IPR009078; Ferritin/RR like.
DR InterPro; IPR000358; Ribonucleotide.
DR Pfam; PF0268; Ribonucleotide.
DR PROSITE; PS00368; RIBORED_SMALL; 1.
KW Hypothetical protein.
SQ SEQUENCE 297 AA; 34498 MW; 2B522F2CC158A02 CRC64;

Alignment Scores:
Pred. No.: 6.61e-126 Length: 297
Score: 1534.00 Matches: 296
Percent Similarity: 99.66% Conservative: 0
Best Local Similarity: 99.66% Mismatches: 1
Query Match: 83.10% Indels: 0
DB: 2 Gaps: 0

US-10-698-228-2 (1-1053) x Q6AI41 (1-297)

QY 163 ATGTATAAACAGGACAGGCTTCTTCTGGACAGAGAGTCTGACTTATCAAAGGAT 222
DB 1 MetTyrLysGlnAlaGlnAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAsp 20
QY 223 CTCCTCATCTGGAACAAGCTTTAAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTAGCC 282
DB 21 LeuProHisTrpAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAla 40
QY 283 TTTTTCGAGCAGGTGGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAG 342
DB 41 PhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGlu 60
QY 343 GTGCAGGTTCCAGAGGCTCGCTGTTTCTATGCTTTTCAAATTCATCCAGATGTTTCAC 402
DB 61 ValGlnValProGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHis 80
QY 403 TCAGAGATGTACAGTTTCTGATAGACACTTACATCAGAGATCCCCAAGAAAAAGGAATTT 462
DB 81 SerGluMetTyrSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPhe 100
QY 463 TTATTTAATGCAATTGAAACCATGCCCTTATGTTAAGAAAAAGCAGATTGGCGCTTGGCA 522
DB 101 LeuPheAsnAlaIleGluThrMetProTyrValLysLysAlaAspTrpAlaLeuArg 120
QY 523 TGGATAGCAGATAGAAATCTACTTTTCGGGAAACAGTGGTGGCTTCTCTCTCTGTAGAA 582
DB 121 TrpIleAlaAspArgLysSerThrPheGlyGluArgValAlaPheAlaAlaValGlu 140
QY 583 GGAGTTTTTCTTCTCAGGATCTTTTGTGCTATATTCTGCTTAAAGAGAGAGGTCTTATG 642

```
Db 141 GlyValPhePheSerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMet 160
QY CCAGACTCACCTTTTCCAAATGAACTCATCAGCAGAGATGAAGGACTTCACCTGTGACTTT 702
Db 161 ProGlyLeuThrPheSerAsnGluLeuIleSerArgSpGluGlyLeuHisCysAspPhe 180
QY GCTTCCTGATGTTCCAACTACTAGTAATAATAGCCCTTCAGAGAAGAGGTCAGGAGATC 762
Db 181 AlaCysLeuMetPheGlnTyrLeuValAsnLysProGluGluArgValArgGluIle 200
QY ATTGTTGATGCTGTCAAATTCAGCAGAGAGTTCCTTAAACAGAACCCCTGCGAGTTCGCTC 822
Db 201 IleValAspAlaValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeu 220
QY ATTGCAATGAATGCTATTTGATGAACAGTACATTTGATGTTGATGTTGATGTTGATGTT 882
Db 221 IleGlyMetAsnCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeu 240
QY GTGGAACCTGGATTCCTCAAGGCTTTTCAGGCAGAGAAATCCTTTTGATTTTATGAAAC 942
Db 241 ValGluLeuGlyPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsn 260
QY ATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTT 1002
Db 261 IleSerLeuGluGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPhe 280
QY GCAGTTATGGCAGAAACACACAGATACGTCCTTCACCTGGGATGCAGATTTT 1053
Db 281 AlaValMetAlaGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 297
```

RESULT 5

```
Q75PY9 PRELIMINARY; PRT; 299 AA.
ID Q75PY9
AC Q75PY9;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DE P53-inducible ribonucleotide reductase small subunit 2 long form.
GN Name=p53R2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ugai H., Yokoyama K.K.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB163437; BAD11774.1; -.
DR GO; GO:0004748; P:ribonucleoside-diphosphate reductase activity; IEA.
DR GO; GO:0009186; P:deoxyribonucleoside diphosphate metabolism; IEA.
DR InterPro; IPR009078; F:feritin/IR-like.
DR InterPro; IPR000358; Ribonuc redctse.
DR Pfam; PF00268; Ribonuc red sm; 1.
DR PROSITE; PS00368; RIBORED_SMALL; 1.
SQ SEQUENCE 299 AA; 34528 MW; 4705C44389EB689B CRC64;
```

Alignment Scores:

```
Pred. No.: 3.78e-124 Length: 299
Score: 1514.00 Matches: 299
Percent Similarity: 85.19% Conservative: 0
Best Local Similarity: 85.19% Mismatches: 0
Query Match: 82.02% Indels: 52
Db: 2 Gaps: 1
```

US-10-698-228-2 (1-1053) x Q75PY9 (1-299)

```
QY 1 ATGGCGACCCGGAAGCCGAGCGCGCGGCTGATCAGGATGAGAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAsp----- 15
QY 61 GACACCAACGAAGTGAATAAAGTCAAATGAAGAGCCACTCTTAAGAAAGAGTTCTTCGC 120
```

```
Db 15 ----- 15
QY 121 CGGTTTGTTCATCTTTCCAAATCCAGTACCCTGATATTGGAAATGTATAAACAGGCACAG 180
Db 15 ----- 15
QY 181 GCTTCCTTCTGGACAGCAGAGAGGTCGACTTATCAAGGATCTCCTCCTCACTGGAACAAG 240
Db 16 -----GluValAspLeuSerLysAspLeuProHisTrpAsnLys 28
QY 241 CTTAAAGCAGATCAGAGTACTTCTCATCTCTCATCTTACCTTTTTCAGCCATTTTCAGCCAGCAT 300
Db 29 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaIleSerAsp 48
QY 301 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
Db 49 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 68
QY 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAATGTTTCACCTCAGAGATGTACAGTTTG 420
Db 69 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 88
QY 421 CTGATAGACATTCATCAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 89 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 108
QY 481 ACCATGCCCTATGTTAAGAAAGCAGATTTGGCCCTTCGATCGATGATGATGATGATGATA 540
Db 109 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 128
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAAGGAGTTTCTTCTCAGGA 600
Db 129 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 148
QY 601 TCTTTTGTCTATATCTGGCTAAAGAGAGAGGCTTATGCGAGAGATCATTTGATGCTGTCAA 780
Db 149 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 168
QY 661 AATGAACTCATCAGCAGAGATGAGGACTTCACCTGTGACTTTGCTGCTGCTGCTGCTGCT 720
Db 169 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 188
QY 721 TACTTAGTAAATAGGCTTCAGAAAGAGGCTCAGGAGAGATCATTTGATGCTGTCAA 780
Db 189 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 208
QY 781 ATTGACGAGGAGTTCCTTAAACAGAACGCTTCCAGTTGGCTCATTTGGAATGAATTCATT 840
Db 209 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 228
QY 841 TTGATCAAAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTTGTGGAACTTGGATTTCTCA 900
Db 229 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 248
QY 901 AAGGTTTTTTCAGGCAGAAAATCCTTTTGTATTTTATGAAACATTTCTTTTAGAAGGAAAA 960
Db 249 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 268
QY 961 ACAAAATTTCTTTCAGAAAACAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020
Db 269 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 288
QY 1021 ACAGATAACGTTCTTCACCTTGGATGCAGATTTT 1053
Db 289 ThrAspAsnValPheThrLeuAspAlaAspPhe 299
RESULT 6
Q86YE3 PRELIMINARY; PRT; 285 AA.
AC Q86YE3;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
```


OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87089677; PubMed=3025593;
RA Thelander L., Berg P.;
RT "Isolation and characterization of expressible cDNA clones encoding
RL the M1 and M2 subunits of mouse ribonucleotide reductase.";
RN Mol. Cell. Biol. 6:3433-3442(1986).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90060004; PubMed=2684652;
RA Thelander M., Thelander L.;
RT "Molecular cloning and expression of the functional gene encoding the
RL M2 subunit of mouse ribonucleotide reductase: a new dominant marker
RN gene.";
RN EMBO J. 8:2475-2479(1989).
[3]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 65-352.
RX MEDLINE=97030707; PubMed=8876648; DOI=10.1006/jmbi.1996.0546;
RA Kauppi B., Nielsen B.B., Ramaswamy S., Larsen I.K., Thelander M.,
RA Thelander L., Eklund H.;
RT "The three-dimensional structure of mammalian ribonucleotide reductase
RL protein R2 reveals a more-accessible iron-radical site than
RT Escherichia coli R2.";
RN J. Mol. Biol. 262:706-720(1996).
[4]
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 65-352.
RX MEDLINE=22206532; PubMed=12087093; DOI=10.1074/jbc.M203358200;
RA Strand K.R., Karlsson S., Andersson K.K.;
RT "Cobalt substitution of mouse R2 ribonucleotide reductase as a model
RL for the reactive diferrous state. Spectroscopic and structural
RT evidence for a ferromagnetically coupled dinuclear cobalt cluster.";
RN J. Biol. Chem. 277:34229-34238(2002).
[5]
RP STRUCTURE BY NMR OF 384-390.
RX MEDLINE=96069780; PubMed=7583667;
RA Fisher A.L., Laub P.B., Cooperman B.S.;
RT "NMR structure of an inhibitory R2 C-terminal peptide bound to mouse
RL ribonucleotide reductase R1 subunit.";
RN Nat. Struct. Biol. 2:951-955(1995).
CC -!- FUNCTION: Provides the precursors necessary for DNA synthesis.
CC -!- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate +
CC thioredoxin disulfide + H(2)O = ribonucleoside diphosphate +
CC thioredoxin.
CC -!- COPACTOR: Binds 2 iron ions per subunit.
CC -!- PATHWAY: DNA replication pathway; first step.
CC -!- SUBUNIT: Heterodimer of a large (M1) and a small chain (M2).
CC -!- MISCELLANEOUS: Two distinct regulatory sites have been defined:
CC the specificity site, which controls substrate specificity, and
CC the activity site which regulates overall catalytic activity. A
CC substrate-binding catalytic site, located on M1, is formed only in
CC the presence of the second subunit M2.
CC -!- SIMILARITY: Belongs to the ribonucleoside diphosphate reductase
CC small chain family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; M14223; AAA40062.1; -;
DR EMBL; X15666; CAA33707.1; -;
DR PIR; S06735; S06735.
DR PDB; 1AFT; NMR; @=383-390.
DR PDB; 1H0N; X-ray; A=1-390.
DR PDB; 1H0O; X-ray; A=1-390.
DR PDB; 1XSM; X-ray; @=1-390.
DR MGD; MGI:98181; Rrm2.
DR GO; GO:0005515; F:protein binding; IPI.

DR GO; GO:0009262; P:deoxyribonucleotide metabolism; IDA.
DR InterPro; IPR009078; Ferritin/RR like.
DR Pfam; PF00268; Ribonuc_redctse.
DR PROSITE; PS00368; RIBORED_SMALL; 1.
KW 3D-structure; DNA replication; Iron; Metal-binding; Oxidoreductase.
FT METAL 139 139 Iron 1.
FT METAL 170 170 Iron 1 and 2.
FT METAL 173 173 Iron 1.
FT METAL 233 233 Iron 2.
FT METAL 267 267 Iron 1 and 2.
FT METAL 270 270 Iron 2.
FT ACT SITE 177 177 By similarity.
FT HELIX 68 70
FT TURN 72 74
FT HELIX 89 99
FT TURN 100 101
FT HELIX 105 107
FT TURN 112 113
FT HELIX 114 118
FT TURN 119 119
FT HELIX 122 148
FT TURN 149 149
FT HELIX 150 153
FT HELIX 157 184
FT HELIX 188 195
FT TURN 196 196
FT HELIX 197 200
FT HELIX 202 204
FT TURN 205 216
FT TURN 218 219
FT HELIX 222 235
FT TURN 236 237
FT HELIX 238 249
FT TURN 250 251
FT HELIX 254 279
FT TURN 280 280
FT HELIX 287 306
FT TURN 307 308
FT HELIX 311 314
FT TURN 315 315
FT HELIX 318 335
FT TURN 336 337
FT HELIX 349 351
SQ SEQUENCE 390 AA; 45095 MW; AC7ACC4FAP8A4A2F CRC64;

Alignment Scores:
Pred. No.: 6.77e-118 Length: 390
Score: 1443.00 Matches: 269
Percent Similarity: 90.62% Conservative: 40
Best Local Similarity: 78.89% Mismatches: 26
Query Match: 78.17% Indels: 6
DB: 1 Gaps: 1

US-10-698-228-2 (1-1053) x RIR2_MOUSE (1-390)
QY 49 AGATCATCTTCAGACACCAACGAAAGTGAATAAGTCA-----AAT 90
Db 50 ArgilepheGlnAspSerAlaGluLeuGluSerLysAlaProThrAsnProSerValGlu 69
QY 91 GAAGAGCCACTCTTAAGAAAGAGTTCTCGCGGTTTGTCTATCTTTTCAATCCAGTACCT 150
Db 70 AspGluProLeuLeuArgGluAsnProArgPheValValPheProIleGluTyrHis 89
QY 151 GATATTGGAAATGTATAACAGGCACAGGCTTCCTTCTGGACAGCAGAGAGTCCGAC 210
Db 90 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluValAsp 109
QY 211 TTATCAAGAGGATCTCCCTCACTGGAAACAGCTTAAAGCAGATGAGAAGTACTTCTCT 270
Db 110 LeuSerLysAspIleGlnHisTrpGluAlaLeuLysProAspGluArgHisPheIleSer 129
QY 271 CACATCTTAGCCTTTTGTGACCCAGTGTGAATTTGAAATTTTGTGTGGAGCC 330

```
Db 130 HisValLeuAlaPhePheAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 149
QY TTTAGTCAGGAGTCCAGGTCGCTGTTCTATGCGTTCCTCAATTCCTCATC 390
Db 150 PheSerGlnGluValGlnValThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 169
QY GAGATGTTCACTCAGAGATGATACAGTTTGTGTAGACACTTACATCAGAGATCCCAAG 450
Db 170 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 189
QY 451 AAAAGGGAATTTTATTATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 510
Db 190 GluArgGluTyrLeuPheAsnAlaIleGluThrMetProCysValLysLysLysAlaAsp 209
QY 511 TGGGCTTGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 570
Db 210 TrpAlaLeuArgTrrPleGlyAspLysGluAlaThrTyrGlyGluArgValValAlaPhe 229
QY 571 GCTGCTGATGAGGAGTTCCTTCTCAGGATCTTTTGTGCTATATCTTGGCTAAAGAG 630
Db 230 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 249
QY 631 AGAGTCTTATGCGAGACTCCTTTTCCATGAACTCATCAGCAGAGATGAAGACTT 690
Db 250 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 269
QY 691 CACTGTGACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 750
Db 270 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProAlaGluGlnArg 289
QY 751 GTCAGGAGATCATTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 810
Db 290 ValArgGluIleIleThrAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 309
QY 811 CCAGTGTGCTCATTTGGAATGAAATTTGATGAAACAGTACATTCAGTTTGTAGCT 870
Db 310 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 329
QY 871 GACAGATTCATTTGGAATCTGGAATCTGGAATCTGGAATCTGGAATCTGGAATCTGGA 930
Db 330 AspArgLeuMetLeuGluLeuGlyPheAsnLysIlePheArgValGluAsnProPheAsp 349
QY 931 TTTATGAAACATTTCTTATAGAGGAAACAAATTTCTTGTAGAAACAGTTCAGAG 990
Db 350 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPheGluLysArgValGlyGlu 369
QY 991 TATCAGCGTTTGTGAGTATGCGAGAAACACAGATAACGCTTCCACCTTGGATGCAGAT 1050
Db 370 TyrGlnArgMetGlyValMetSerAsnSerThrGluAsnSerPheThrLeuAspAlaAsp 389
QY 1051 TTT 1053
Db 390 Phe 390
RESULT 9
Q68EP1 PRELIMINARY; PRT; 386 AA.
AC Q68EP1;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8364;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
```

```
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zordan B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards D.K., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Gerhard D.S.;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC080161; AAH80161.1; -.
DR InterPro; IPR009078; Ferritin/RR like.
DR Pfam; PF00268; Ribonuc_red_sm; 1.
DR PROSITE; PS00368; RIBORED_SMALL; 1.
KW Hypothetical protein.
SQ SEQUENCE 386 AA; 44540 MW; 277DF859B6053824 CRC64;
Alignment Scores:
Pred. No.: 1,52e-117 Length: 386
Score: 1439.00 Matches: 268
Percent Similarity: 89.50% Conservative: 39
Best Local Similarity: 78.13% Mismatches: 36
Query Match: 77.95% Indels: 0
DB: 2 Gaps: 0
US-10-698-228-2 (1-1053) x Q68EP1 (1-386)
QY 25 GCGCGCGCGCTGGATCAGGATGAGATCATCTTCAGACACCAACGAAAGTGAATAAAG 84
Db 44 AlaArgAsnIlePheGlnGluAlaGluThrLysSerLysAlaProLysAspProArg 63
QY 85 TCAATGTAAGAGCCACTCTCTAAGAAAGAGTCTTCGCGGTTTGTCTATCTTCCAAATCCAG 144
Db 64 IleGlnAspGluProLeuLeuLysAspAsnProHisArgPheValIlePheProIleGln 83
QY 145 TACCTGTATTTGGAAATGTATAACAGGACAGGCTTCCTCTCGACAGCAGAGAGAG 204
Db 84 TyrHisAspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGlu 103
QY 205 GTCGACTTATCAAGGATCTCCCTCACTGCAACCAAGCTTAAAGCAGATGAGAACTACTTC 264
Db 104 ValAspLeuSerLysAspLeuArgHisTrpGluSerLeuLysAlaGluGluLysTyrPhe 123
QY 265 ATCTCTCATCTTAGCTTTTTCAGCCAGTGTGGAATTTGTAATGAAATTTGGTG 324
Db 124 IleSerHisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuVal 143
QY 325 GAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCTCGCTGTTCTATGCTTCAATTC 384
Db 144 GluArgPheSerLysGluValGlnValThrGluAlaArgCysPheTyrGlyPheGlnIle 163
QY 385 CTCATCGAAGATGTTCACTCAGAGATGTACAGTTTGTCTGATACACTTTCATCATCAGAGAT 444
Db 164 AlaMetGluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAsp 183
QY 445 CCAGAGAAAGGGAATTTTATTTATGCAATTTGAAACCAATGCCTATGTTAAGAAAAA 504
Db 184 ProLysGluArgGluTyrLeuPheAsnAlaIleGluThrLeuProCysValLysLysLys 203
```


251 GlyLeuThrPheSerAsnGluLeuLeuSerArgAspGluGlyLeuHisCysAspPheAla 270
706 TGCCTGATGTTCCATACTAGTAATAAGCTTCAGAGAAAGGCTCAGGAGATCATT 765
271 CysLeuMetPheLysHisLeuLeuArgLysProSerGluGluArgValValGluLeuLeu 290
766 GTTGATGCTGTCAAAATTCAGCAGGAGTTCCTTAAACAGAGCCCTCCAGCTGGCTCATT 825
291 ThrAspAlaValGlnLeuGluGlnGluPheLeuThrGluAlaLeuProValAsnLeuLeu 310
826 GGAATGAATTCATTTGATGAACAGTACATAGTGTGTTAGCTGACAGATTCATTGTG 885
311 GlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuLeu 330
886 GAACCTTGATTCCTCAAGGTTTTCAGGCAGAAATCCTTTGATTTTATGGAACATTT 945
331 GluLeuGlyPheLysLysValPheLysAlaThrAsnProPheAspPheMetGluAsnLeu 350
946 TCTTTAGAGGAAACAAATTTCTTTGAGAACAGATTTTCAGAGTATCAGCGTTTTCGA 1005
351 SerLeuGluGlyLysThrAsnPhePheGluLysLysValGlyGluTyrGlnLysMetGly 370
1006 GTTATGGCAGAAACACAGATAACCTCTTCACCTTGGATGAGATTTT 1053
371 ValMetSerLysAlaLysAspAsnThrPheThrLeuAspAlaAspPhe 386

RESULT 12
RIR2_BRARE ID RIR2_BRARE STANDARD; PRT; 386 AA.
AC P79733;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Ribonucleoside-diphosphate reductase M2 chain (EC 1.17.4.1)
DE (Ribonucleotide reductase protein R2 class I).
GN Name=rrm2; Synonyms=r2;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
[1]
SEQUENCE FROM N.A.
RX MEDLINE=97137859; PubMed=8983196;
RA Mathews C.Z., Sjoeborg B.-M., Karlsson M.;
RT "Cloning and sequencing of cDNAs encoding ribonucleotide reductase
from zebrafish Danio rerio."
RL Mol. Mar. Biol. Biotechnol. 5:284-287(1996).
CC -!- FUNCTION: Provides the precursors necessary for DNA synthesis.
CC -!- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate +
CC thioredoxin disulfide + H(2)O = ribonucleoside diphosphate +
CC thioredoxin.
CC -!- COFACTOR: Binds 2 iron ions per subunit (By similarity).
CC -!- PATHWAY: DNA replication pathway; first step.
CC -!- SUBUNIT: Heterodimer of a large (M1) and a small chain (M2).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the ribonucleoside diphosphate reductase
CC small chain family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; U57965; AAB37103.1; -;
DR HSPB; P11157; IYXSN.
DR ZFIN; ZDB-GENE-990415-25; rrm2.
DR InterPro; IPR009078; Ferritin/RR like.
DR InterPro; IPR000358; Ribonucl_redctse.

DR Pfam; PF00268; Ribonuc_red_sm; 1.
KW PROSITE; PS00368; RIBORED_SMALL; 1.
FT METAL 135 135 Iron 1 (By similarity).
FT METAL 166 166 Iron 1 and 2 (By similarity).
FT METAL 169 169 Iron 1 (By similarity).
FT METAL 229 229 Iron 2 (By similarity).
FT METAL 263 263 Iron 2 (By similarity).
FT METAL 266 266 Iron 2 (By similarity).
FT ACT_SITE 173 173 By similarity.
SQ SEQUENCE 386 AA; 44594 MW; C168846FB57F9F4E CRC64;

Alignment Scores:
Pred. No.: 2,33e-116 Length: 386
Score: 1425.50 Matches: 268
Percent Similarity: 89.41% Conservative: 36
Best Local Similarity: 78.82% Mismatches: 31
Query Match: 77.22% Indels: 5
DB: 1 Gaps: 1

US-10-698-228-2 (1-1053) x RIR2_BRARE (1-386)
QY 49 AGATCATCTTCAGACACCAACGAAAGTGAATAAAGTCA-----AATGAA 93
DB 47 ArgLysIlePheAspGluSerGluGlyGlnSerLysAlaLysLysGlyAlaValGluGlu 66
QY 94 GAGCCACTCTAAGAAAGAGTTCTCGCCGTTTGTCTATCTTTCCAATCCAGTACCCCTGAT 153
DB 67 GluProLeuLeuLysGluAsnProHisArgPheValIlePheProIleGlnTyrHisAsp 86
QY 154 ATTTGGAAATGTATAACAGGCACACAGGCTCTCTTCTGGACAGCAGAAAGGTCGACTTA 213
DB 87 IleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAspLeu 106
QY 214 TCAAGAGATCTCCCTCAGTGGAAACAGCTTAAGACAGATGAGAAAGTACTTCACTCTCAC 273
DB 107 SerLysAspLeuGlnHisTrpAspSerLeuLysAspGluGluArgTyrPheIleSerHis 126
QY 274 ATCTTAGCTCTTTTTCAGCCAGTGTGAATTTGTAATGAAATTTGGTGGAGCGCTTT 333
DB 127 ValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArgPhe 146
QY 334 AGTCAGGAGGTGCAGGCTCCAGAGGCTCGCTGTTCTATGGCTTTCAAATTTCTCATCGAG 393
DB 147 ThrGlnGluValGlnValThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMetGlu 166
QY 394 AATGTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAGATCCCAAGAA 453
DB 167 AsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspSerLysGlu 186
QY 454 AGGGAATTTTATTAAATGCAATTGAAACCATGCCCTATGTTAAAGAAAAAGCAGATTGG 513
DB 187 ArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysLysAlaAspTrp 206
QY 514 GCCTTTCGATGATAGCAGATCAATAAATCTACTTTTGGGAAAGAGTGGTGGCTTTGCT 573
DB 207 AlaLeuAsnTrpIleGlyAspLysAsnAlaArgTyrGlyGluArgValValAlaPheAla 226
QY 574 GCTGTAGAAAGAGATTTTCTCTCAGGATCTTTTGTGCTATATTTCTGGCTTAAGAAAGAG 633
DB 227 AlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLysArg 246
QY 634 GGTCTATGCCAGGACTCACTTTTCCATGAATCACTCATCAGCAGAGATGAAGACTTCAC 693
DB 247 GlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHis 266
QY 694 TGTGACTTTTGTGCTGATGTTTCCAATACTTAGTAAATTAAGCCTTCAGAAAGAGGGTC 753
DB 267 CysAspPheAlaCysLeuMetPheLysHisLeuIleAsnLysProSerGluGluThrVal 286
QY 754 AGGAGATCATTGTTGATGCTGTCAAAATGAGCAGAGAGTTTAAACAGAAAGCTTCCCA 813
DB 287 LysLysIleIleMetAsnAlaValGluIleGluGlnGluPheLeuThrAspAlaLeuPro 306

QY 814 GTTGCCCTCATTTGGAATGAATTCATTTTGTGATGAACAGTACATTGAGTTTGTAGCTGAC 873
FT METAL 170 170 Iron 1 and 2 (By similarity).
FT METAL 173 173 Iron 1 (By similarity).
FT METAL 233 233 Iron 2 (By similarity).
Db 307 VallysLeuileGlyMetAsnCysAspLeuMetLysGlnTyrIleGluPheValAlaAsp 326
FT METAL 267 267 Iron 2 (By similarity).
FT METAL 270 270 Iron 2 (By similarity).
QY 874 AGATTACTTGTGGAACCTGGATTCTCAAGGTTTTCAGCGAGAAATCCTTTGATTTT 933
FT ACT_SITE 177 177 By similarity.
Db 327 ArgLeuLeuGluLeuGlyPheAspLysValTyArgValGluAsnProPheAspPhe 346
QY 934 ATGGAATAACATTTCTTTAGAAGGAAACAAATTTCTTTGAGAAACGAGTTTCAGAGTAT 993
Db 347 MetGluAsnIleSerLeuGluGlyysThrAsnPheGluysbargValGlyGluTy 366
QY 994 CAGCGTTTTTGCAGTTATGCGAGAAACCCAGATAACGCTTCCACCTTGTGATGACAGATTTT 1053
Db 367 GlnArgMetGlyValMetSerGlyThrThrAspAsnThrPheThrLeuAspAlaAspPhe 386

RESULT 13
RIR2_MESAU
ID RIR2_MESAU STANDARD; PRT; 386 AA.
AC Q60561;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Ribonucleoside-diphosphate reductase M2 chain (EC 1.17.4.1)
DE (Ribonucleotide reductase small chain).
GN Name=RM2;
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]_TaxID=10036;
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=93042000; PubMed=1394717; DOI=10.1016/0167-4781(92)90151-0;
RA Chaudhuri M.M., Tonin P.N., Srinivasan P.R.;
RT "cDNA sequence of the small subunit of the hamster ribonucleotide reductase";
RL Biochim. Biophys. Acta 1171:117-121(1992).
CC -!- FUNCTION: Provides the precursors necessary for DNA synthesis.
CC -!- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate +
CC thioresoxin disulfide + H(2)O = ribonucleoside diphosphate +
CC thioredoxin.
CC -!- COFACTOR: Binds 2 iron ions per subunit (By similarity).
CC -!- PATHWAY: DNA replication pathway; first step.
CC -!- SUBUNIT: Heterodimer of a large (M1) and a small chain (M2).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- MISCELLANEOUS: Two distinct regulatory sites have been defined:
CC the specificity site, which controls substrate specificity, and
CC the activity site which regulates overall catalytic activity. A
CC substrate-binding catalytic site, located on M1, is formed only in
CC the presence of the second subunit M2.
CC -!- SIMILARITY: Belongs to the ribonucleoside diphosphate reductase
CC small chain family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X68127; CAA48232.1; -
CC PIR; S27153; S27153.
DR HSSP; P11157; 1XSM.
DR InterPro; IPR009078; Ferritin/RR like.
DR InterPro; IPR000358; Ribonucleotide reductase.
DR Pfam; PF00268; Ribonucleotide reductase sm; 1.
DR PROSITE; PS00368; RIBORED_SMALL; 1.
KW DNA replication; Iron; Metal-binding; Oxidoreductase.
FT METAL 139 139 Iron 1 (By similarity).

FT METAL 170 170 Iron 1 and 2 (By similarity).
FT METAL 173 173 Iron 1 (By similarity).
FT METAL 233 233 Iron 2 (By similarity).
Db 307 VallysLeuileGlyMetAsnCysAspLeuMetLysGlnTyrIleGluPheValAlaAsp 326
FT METAL 267 267 Iron 2 (By similarity).
FT METAL 270 270 Iron 2 (By similarity).
QY 874 AGATTACTTGTGGAACCTGGATTCTCAAGGTTTTCAGCGAGAAATCCTTTGATTTT 933
FT ACT_SITE 177 177 By similarity.
Db 327 ArgLeuLeuGluLeuGlyPheAspLysValTyArgValGluAsnProPheAspPhe 346
QY 934 ATGGAATAACATTTCTTTAGAAGGAAACAAATTTCTTTGAGAAACGAGTTTCAGAGTAT 993
Db 347 MetGluAsnIleSerLeuGluGlyysThrAsnPheGluysbargValGlyGluTy 366
QY 994 CAGCGTTTTTGCAGTTATGCGAGAAACCCAGATAACGCTTCCACCTTGTGATGACAGATTTT 1053
Db 367 GlnArgMetGlyValMetSerGlyThrThrAspAsnThrPheThrLeuAspAlaAspPhe 386

Alignment Scores:
Pred. No.: 5,79e-116 Length: 386
Score: 1421.00 Matches: 266
Percent Similarity: 91.57% Conservative: 38
Best Local Similarity: 80.12% Mismatches: 20
Query Match: 76.98% Indels: 8
DB: 1 Gaps: 2

US-10-698-228-2 (1-1053) x RIR2_MESAU (1-386)

QY 70 GAAAGTGAATAAAGTCAAAAT-----GAAGAGCCACTCTCTAAGAAAGAGTTCT 117
Db 59 GluSerLysValSerThrAsnProSerValGluAspGluProLeuLeuArgGluAsnPro 78
QY 118 CGCGGTTTGTCTATCTTCCAAATCCAGTACCCCTGATATTTGGAAAATGTATAACACGCCA 177
Db 79 ArgArgPheValValPheProIleGluTyHisAspIleTrpLysMetTyLysLysAla 98
QY 178 CAGGCTTCTCTTGGACAGCAGAGAGGTGCTGACTTATCAAGGATCTCCCTCACTGGAAC 237
Db 99 GluAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspIleGlnHisTrpGlu 118
QY 238 AAGCTTTAAAGCAGATGAGAACTACTTCACTCTCTCACTTTAGCCTTTTTCAGCCAGT 297
Db 119 AlaLeuLysProAspGluArgHisPheIleSerHisValLeuAlaPhePheAlaAsp 138
QY 298 GATGGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGTCCAGAG 357
Db 139 AspGlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValTrpGlu 158
QY 358 GCTCGCTGTTTCTATGCTTCAAAATTTCTCATCGAATGTTCTCACTCAGAGATGACAGT 417
Db 159 AlaArgCysPheTyGlyPheGlnIleAlaMetGluAsnIleHisSerGluMetTySer 178
QY 418 TTGCTCATAGACACTTACATCAGATCCCAAGAAAGGAAATTTTATTTAATGCAATT 477
Db 179 LeuLeuIleAspThrTyIleLysAspSerLysGluArgGluTyLeuPheAsnAlaIle 198
QY 478 GAAACATGCCCTATGTTAAGAAAAAGCAGATGGGCTTGGCATGGATGATAGCAGATAGA 537
Db 199 GluThrMetProCysValLysLysAlaAspTrpAlaLeuArgTrpIleGlyAspLys 218
QY 538 AAATCTACTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAGAGTTCCTCTCA 597
Db 219 GluAlaThrTyGlyGluArgValValAlaPheAlaValGluGlyIlePhePheSer 238
QY 598 GGATCTTTTGTGCTATATTTCTGCTTAAAGAGAGAGTCTTATGCCAGACTCACATTTT 657
Db 239 GlySerPheAlaSerIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuPhe 258
QY 658 TCAATGAATCACTCAGCAGATGAAGGACTTCACTGTGACTTTGCTGCTGATGTTTC 717
Db 259 SerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPhe 278
QY 718 CAATCTTGTAGTAATAAGCTTTCAGAAAGAGGTGAGGAGATCATTTGTGATGCTGTC 777
Db 279 LysHisLeuValHisLysProSerGluGlnArgValGlnGluIleIleThrAsnAlaVal 298
QY 778 AAAATGACGAGGAGTTTAAACAGAAAGCTTCCAGTTGGCTCATTTGGAATGAATGTC 837
Db 299 ArgIleGluGlnGluPheLeuThrGluAlaLeuProValLysLeuIleGlyMetAsnCys 318
QY 838 ATTTTGTGATCAACAGTACATTGATTTGTAGCTGACAGATTAATTCTGTGGAACTTGGATTC 897
Db 319 ThrLeuMetLysGlnTyIleGluPheValAlaAspArgLeuMetLeuGluLeuGlyPhe 338

US-10-598-228-2 (1-1053) x Q6DI44 (1-386)

Qy	49	AGATCATCTTCAGACACCAACGAAAGTGAAATAAAGTCA-----AATGAA	93
Db	47	ArgLysIlePheAspGluSerGluGlyGlnSerLysAlaLysLysGlyAlaValGluGlu	66
Qy	94	GAGCCACTCTAAGAAAGAGTTCTCCCGGTGTTGTCTATCTTTTCCAATCCAGTACCCTGAT	153
Db	67	GluProLeuLeuLysGluAsnProHisArgPheValIlePheProIleGlnTyrHisAsp	86
Qy	154	ATTTTGGAATGTATAAACAGGCACAGGCTTCTCTGGACACAGAGAGGTCTGACTTA	213
Db	87	IleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAspLeu	106
Qy	214	TCAAGGATCTCCCTCAGTGAACAAAGCTTAAAGCAGATGAGAGTACTTCTCTCTCAC	273
Db	107	SerLysAspLeuGlnHisTrpAspSerLeuLysAspGluGluArgTyrPheIleSerHis	126
Qy	274	ATCTTAGCCTTTTTCGAGCCAGTGATGGAAATGTAAATGAAAAATTTGTGTGGAGCGCTTT	333
Db	127	ValLeuAlaPhePheAlaLaaSerAspGlyIleValAsnGluAsnLeuValGluArgPhe	146
Qy	334	AGTCAGGAGGTTCAGGTTTCCAGAGGTCGCTGTTTCTATGGCTTCCAAATTCATCTCGAG	393
Db	147	ThrGlnGluValGlnValThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMetGlu	166
Qy	394	AATGTTTCACATCAGAGATGTACAGTTTGCTGATAGACACTTACATCAGAGATCCCAAGAAA	453
Db	167	AsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspSerLysGlu	186
Qy	454	AGGGAATTTTATTTAATGCAATTTGAAACCATCCCTATGTTTAAGAAAAAAGCAGATTGG	513
Db	187	ArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysLysAlaAspTrp	206
Qy	514	GCCTTGCATGGATAGCAGATAGAAAATCTACTTTTGGGGAAGAGTGGTGGCTTTGCT	573
Db	207	AlaLeuAsnTrpIleGlyAspLysAsnAlaArgTyrGlyGluArgValValAlaPheAla	226
Qy	574	GCTGTAGAAGAGTTTTCTCTCAGGATCTTTTGGCTGTATATCTGGCTTAAAGAGAGA	633
Db	227	AlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLysArg	246
Qy	634	GGTCTTATGCCAGGACTCACTTTTCCAATGAACATCATCAGCAGAGATGAAGACTTCAC	693
Db	247	GlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHis	266
Qy	694	TGTGACTTCTTGCCTGATGTTTCCAATCACTTAGTAATAAAGCCTTCAGNAGAAAGGTC	753
Db	267	CysAspPheAlaCysLeuMetPheLysHisLeuIleSerLysProSerGluGluThrVal	286
Qy	754	AGGAGATCAATTGTTGATGCTGTCAAAATTTGACGAGGATTTTAAACAGAAAGCTTGCCA	813
Db	287	LysLysIleIleMetAsnAlaValGluIleGluGlnGluPheLeuThrAspAlaLeuPro	306
Qy	814	GTTGGCCTCATTTGGAATGAATTCATTTTGAATGAAACAGTACATGAGTTTGTAGCTGAC	873
Db	307	ValLysLeuIleGlyMetAsnCysAspLeuMetLysGlnTyrIleGluPheValAlaAsp	326
Qy	874	AGATTACTTGGAACCTTGATCTCAAAGTTTTTCAGGCAGAAAAATCCCTTTGATTTT	933
Db	327	ArgLeuLeuLeuGluLeuGlyPheAspLysValTyrArgValGluAsnProPheAspPhe	346
Qy	934	ATGGAAACATTTCTTTTAGAAGGAAAAACAAATTTCTTTGAGAAACAGGTTTCAGAGTAT	993
Db	347	MetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGluTyr	366
Qy	994	CAGCGTTTTTCAGTTATGCGCAAAACACAGATAAAGCTTCTTCACTTTGGATGAGATTTT	1053
Db	367	GlnArgMetGlyValMetSerGlyThrThrAspAsnThrPheThrLeuAspAlaAspPhe	386

RESULT 15
Q6P876
ID Q6P876 PRELIMINARY: PRT; 378 AA.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 30, 2005, 06:01:36 ; Search time 15.5 Seconds

(without alignments)
10142.641 Million cell updates/sec

Title: US-10-698-228-2

Perfect score: 1846

Sequence: 1 atggggaccggaagcc.....tcaccttgatgcagatttt 1053

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 1027090

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US10698228/runat 26102005 100609 4010/app query.fasta_1.2446
-DB=Issued Patents AA -QFMT=fastan -SUFFIX=n2p.ra1 -MINMATCH=0.1 -LOOPCI=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-USER=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-MODE=US10698228 @Cgn 1 32 @runat 26102005 100609 4010 -NCPUS=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/6C_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1821	98.6	351	4	US-10-019-733-1
2	1454	78.8	389	4	US-10-019-733-5
3	1454	78.8	389	4	US-09-949-016-6016
4	1454	78.8	453	4	US-09-949-016-7896
5	1131	61.3	411	4	US-09-248-796A-18017
6	1072.5	58.1	274	1	US-08-307-499-19
7	1072.5	58.1	274	3	US-09-299-268-19
8	974.5	52.8	322	1	US-08-136-743B-2
9	969.5	52.5	360	4	US-09-248-796A-18018
10	768	41.6	147	3	US-08-905-223-358
11	599	32.4	149	4	US-09-270-767-46236
12	495	26.8	323	4	US-09-902-540-10006

13	381.5	20.7	445	4	US-09-328-352-4714	Sequence 4714, Ap
14	377.5	20.4	476	4	US-09-252-991A-19358	Sequence 19358, A
15	335	18.1	365	4	US-09-198-452A-1059	Sequence 1059, Ap
16	335	18.1	380	4	US-09-438-185A-986	Sequence 986, App
17	299.5	16.3	121	1	US-08-307-499-7	Sequence 7, Appl
18	299.5	16.3	121	3	US-09-299-268-7	Sequence 7, Appl
19	298	16.1	314	3	US-09-080-044-3	Sequence 3, Appl
20	298	16.1	314	3	US-09-531-857A-3	Sequence 3, Appl
21	294	15.9	56	3	US-08-905-223-362	Sequence 362, App
22	220	11.9	382	4	US-09-543-681A-4239	Sequence 4239, Ap
23	203.5	11.0	412	4	US-09-489-039A-8180	Sequence 8180, Ap
24	201.5	10.9	401	4	US-09-540-236-3566	Sequence 3566, Ap
25	200	10.8	325	4	US-09-107-532A-4443	Sequence 4443, Ap
26	195.5	10.6	320	3	US-09-092-437-2	Sequence 2, Appl
27	195.5	10.6	320	4	US-09-583-110-3858	Sequence 3858, Ap
28	195.5	10.6	320	4	US-09-107-433-3136	Sequence 3136, Ap
29	195	10.6	339	4	US-09-543-681A-5069	Sequence 5069, Ap
30	169.5	9.2	323	4	US-09-042-775-2	Sequence 2, Appl
31	167	9.0	343	3	US-09-134-001C-5484	Sequence 5484, Ap
32	163	8.8	314	4	US-09-134-000C-5726	Sequence 5726, Ap
33	161	8.7	311	4	US-09-710-279-1260	Sequence 1260, Ap
34	157	8.5	324	2	US-08-813-940-2	Sequence 2, Appl
35	156	8.5	391	4	US-09-489-039A-10121	Sequence 10121, A
36	153	8.3	104	4	US-09-230-485-5	Sequence 5, Appl
37	138	7.5	321	2	US-08-813-940-6	Sequence 6, Appl
38	127	6.9	52	3	US-08-917-299-21	Sequence 21, Appl
39	127	6.9	52	3	US-09-422-662-21	Sequence 21, Appl
40	127	6.9	52	4	US-09-730-763-21	Sequence 21, Appl
41	127	6.9	52	4	US-09-429-370-21	Sequence 32, Appl
42	103	5.6	555	3	US-09-173-151A-32	Sequence 32, Appl
43	100.5	5.4	382	4	US-09-134-000C-5889	Sequence 5889, Ap
44	99	5.4	233	3	US-08-993-825-2	Sequence 2, Appl
45	98.5	5.3	896	4	US-09-270-767-46130	Sequence 46130, A

ALIGNMENTS

RESULT 1
US-10-019-733-1
; Sequence 1, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619W00P
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 1
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-1

Alignment Scores:
Pred. No.: 4,89e-211 Length: 351
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.65% Indels: 0
DB: 4 Gaps: 0

US-10-698-228-2 (1-1053) x US-10-019-733-1 (1-351)

QY

1 ATGGGCGACCCGGAAGCGCGCGGTGGATCAGATGAGATCATCTTCA 60

```
Db 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAAGGAAAGTAAAGTCAATGAAGAGCCACTCTAAGAAGAGTTCTCCG 120
Db 21 AspThrAsnGluSerGluIleYsSerAsnGluGluProLeuLeuArgLysSerArg 40
QY 121 CGGTTGTGCATCTTTCCCAATCCAGTACCCTGATATTGTGAAAATGTATAAACAAGGCACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCTCTTGACAGCAGAGAGGTCGACTTATCAAAAGGATCTCCCTCAGTGGACAAG 240
Db 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAGCAGATGAGAGTACTTCTCTCTCATCTCTACATCTTACGCTTTTTCGACCCAGTAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY 301 GGAATTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValProGluAla 120
QY 361 CGCTGTTCTATGCGTTTCAAAATCTCATPCGAGAATGTTCTACTCAGAGATGACAGTTTG 420
Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATGAA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAAGAAAAAGCAGATTGGCGCTTCGGATGGATAGCAGATAGAAA 540
Db 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTGTGCTGTGAAGAGAGTTTCTTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValIleGluValPhePheSerGly 200
QY 601 TCTTTTGTGCTATATTCTGGCTTAAGAAGAGAGGTCTTATGCCAGGACTACTTTTTC 660
Db 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACATCATCAGCAGAGATCAAGGACTTCACTGTGACTTGTGCTGTGCTGATGTTCCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAAATTAAGCCTTTCAGAGAAAGGGTCAGGAGATCATTTGTTGATGCTGTCAA 780
Db 241 TyrLeuValAsnLysProSerGluArgValArgGluIleValAspAlaValLys 260
QY 781 ATGAGCAGGAGTTTAAACAGAGCCTTGGCAGTTGGCCTCATTTGGAATGCAATTCAT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAACAGTACATTGAGTTCTAGCTGACAGATTACTTGTGGAACTTGGATCTCA 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTCAGCAGAAAAATCCTTTGATTTTATGGAACAATTTCTTTAGAGAGAAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAAAATTTCTTGAAACAGAGTTTTCAGAGTATCAGCGTTTTCAGATTATGCGAGAAACC 1020
Db 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAAGCTTTCACCTTGAATGCGAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
```

RESULT 2

US-10-019-733-5

```
; Sequence 5, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WO0P
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 5
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-5

Alignment Scores:
Pred. No.: 1,27e-166 Length: 389
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: 4 Gaps: 0

US-10-698-228-2 (1-1053) x US-10-019-733-5 (1-389)
```

```
QY 91 GAAGAGCCACTCTCTAAGAAAGAGTTCTCGCGGGTTTGTCTTCATCTTCCATCCAGTACCCT 150
Db 69 AspGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGluTyrHis 88
QY 151 GATATTGTGAAAATGTATAAACAAGGCACAGCTTCTCTCGACAGCAGAGAGGTCCGAC 210
Db 89 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 108
QY 211 TTATCAAAAGGATCTCCCTCACTCGAACCAAGCTTAAAGCAGATGAGAAAGTACTTCTCTCT 270
Db 109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128
QY 271 CACATCTTAGCCCTTTTGTGACGAGTGTGAATGTAATGAAATTAATGAAATTTGGTGGAGCGC 330
Db 129 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148
QY 331 TTTAGTCAGAGGTGCGAGTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAATCTCATC 390
Db 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
QY 391 GAGAATGTTCACTCAGAGATGTACAGTTTCTGCTAGACACTTACATCAGAGATCCCAAG 450
Db 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
QY 451 AAAAGGGAATTTTATTAATGCAATTTGAAACCAACGATCCCTATGTTAAGAAAAAGCAGAT 510
Db 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 208
QY 511 TGGGCTTTCGAGTGGATAGCAGATAGAAATCTACTTTTGGGGAAGAGTGGTGGCCCTTT 570
Db 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaAlaPhe 228
QY 571 GCTGCTGTGAAGAGGTTTCTCTCAGGATCTTTTCTGCTATATTCTGCTGCTAAAGAG 630
Db 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 248
QY 631 AGAGTCTTATGCGCAGGACTCACTTTTCCAAATGAACATCAGCAGAGATGAAGGACTT 690
Db 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
```

QY 691 CACTGTGACCTTCTGCTGATGTTCCAACTTAGTAAATAGCCTTCAGAGAAAGG 750
 Db 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
 QY 751 GTCAGGAGATCATCTGCTGCTCAAAATGACAGAGCTTTTAAACAGAGCCTTG 810
 Db 289 ValArgGluLeuLeuLeuLeuLeuAlaValArgGluGluGlnGluPheLeuThrGluAlaLeu 308
 QY 811 CCAGTTGGCCTCATTTGGAATGAAATTCATTTGATGAAACAGACATTTGATTTGTAGCT 870
 Db 309 ProValLysLeuLeuGlyMetLysCysThrLeuMetLysGlnTyrLeuGluPheValAla 328
 QY 871 GACAGATTACTTGTGGAATCTGATCTCAAGGTTTTCAGCAGAGAAATCTCTTTGAT 930
 Db 329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348
 QY 931 TTTATGGAACATTTCTTTAGAGAGAAACAAATTTCTTTCAGAGAAACGAGTTTCAGAG 990
 Db 349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368
 QY 991 TATCAGCGTTTTCAGTTATGCGAGAAACACAGATAACGCTCTTCACTTGGATGCAGAT 1050
 Db 369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
 QY 1051 TTT 1053
 Db 389 Phe 389
 RESULT 3
 US-09-949-016-6016
 ; Sequence 6016, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 6016
 ; LENGTH: 389
 ; TYPE: PRT
 ; ORGANISM: Human
 ; US-09-949-016-6016
 Alignment Scores:
 Pred. No.: 1,27e-166 Length: 389
 Score: 1454.00 Matches: 269
 Percent Similarity: 94.08% Conservative: 33
 Best Local Similarity: 83.80% Mismatches: 19
 Query Match: 78.76% Indels: 0
 DB: 4 Gaps: 0
 US-10-698-228-2 (1-1053) x US-09-949-016-6016 (1-389)
 QY 91 GAAGAGCCACTCTAAGAAAGATGTTCTCGCGGTTTGTCTATCTTCCAAATCCAGTACCTT 150
 Db 69 AspGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGluTyrHis 88
 QY 151 GATATTGAAATGTATAAACAGGCACAGGCTTCTTTCGACAGCAGAGAGGTCGAC 210
 Db 89 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluValAsp 108
 QY 211 TTATCAAGAGTCTCCCTCAGTGAACAGCTTAAAGCAGATGAGAGTACTTCACTCT 270

Db 109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128
 QY 271 CACATCTTAGCTTTTTCAGCCAGTATGGAATTTAAATGAAATTTGTGTGAGCCG 330
 Db 129 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148
 QY 331 TTTAGTCAGGAGGTCAGGCTCCAGAGCTCGCTGTTTCTATGCTTTCAAAATTCATC 390
 Db 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
 QY 391 GAGATGTTCTACTCAGAGATGTACAGTTGCTGTGATAGACACTTACATCAGAGATCCCAAG 450
 Db 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
 QY 451 AAAAGGGAATTTTATTAAATGCAATTAACCAATGCCCTATGTTAAAGAAAAACAGAT 510
 Db 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 208
 QY 511 TGGGCTTTCAGATGATAGCAGATAGAAATCTACTTTTGGGGAAGAGTGTGCTGCTTT 570
 Db 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaPhe 228
 QY 571 GCTGCTCTAGAGAGTTTCTTCTCAGGATCTTTTCTGCTGTATATTCTGCTAAAGAG 630
 Db 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 248
 QY 631 AGAGTCTTATGCGAGACTCAGCTTTTCCAAATCACTCACTCAGCAGAGATGAAGACTT 690
 Db 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
 QY 691 CACTGTGACTTTCCTGCTGATGTTCCAAATCACTTAGTAAATAGCCTTCAGAGAAAGG 750
 Db 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
 QY 751 GTCAGGAGATCATTTGTTGATGCTGCTCAAAATTCAGCAGAGAGTTTTAAACAGAGCCTTG 810
 Db 289 ValArgGluLeuIleIleAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 308
 QY 811 CCAGTTGGCCTCATTTGGAATGAAATTCATTTGATGAAACAGTACATGAGTTTGTAGCT 870
 Db 309 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 328
 QY 871 GACAGATTACTTGTGGAATCTGCTCAAGGTTTTCAGGCGAGAAATCCTTTTTCAT 930
 Db 329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348
 QY 931 TTTATGGAACATTTCTTTAGAGAGAAACAAATTTCTTTGAGAAACGAGTTTTCAGAG 990
 Db 349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368
 QY 991 TATCAGCGTTTTCAGTTATGCGAGAAACACAGATAACGCTTCACTTGGATGCAGAT 1050
 Db 369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
 QY 1051 TTT 1053
 Db 389 Phe 389
 RESULT 4
 US-09-949-016-7896
 ; Sequence 7896, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR FILING DATE: 2000-10-03

[illegible]


```

|||||
201 CysAspPheAlaCysLeuMetPheLysHisLeuLeuHisProSerLysGluValIle 220
QY 754 AGGGAGATCATGTTGATGCTGCTCAAAATGAGCAGGAGTTTAAACAGAGCCCTGGCCA 813
Db 221 ThrSerIleIleIleAspAlaValAsnIleGluLysGluPheLeuThrValAlaIlePro 240
QY 814 GTTGGCCTCATTTGGAATGAAATTTGATGAAACAGTACATTTGAGTTGTAGCTGAC 873
Db 241 ValAspLeuIleGlyMetAsnCysCysLeuMetSerGlnIryIleGluPheValAlaAsp 260
QY 874 AGATTACTTTGTGGAACCTTGGATTCTCAAG 903
Db 261 ArgLeuLeuThrGluLeuGlyCysGluLys 270

```

RESULT 7

```

US-09-299-268-19
; Sequence 19, Application US/09299268
; Patent No. 6217882
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; TITLE OF INVENTION: Live Vaccine Vector
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: U.S.A.
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/299,268
; FILING DATE:
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/901,127
; FILING DATE:
; APPLICATION NUMBER: US 07/908,241
; FILING DATE: 1-JUL-1992
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,630
; FILING DATE: 29-JUN-1992
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/342,212
; FILING DATE: 21-APR-1992
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF35.1.FWCC1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 274 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-299-268-19

```

```

Alignment Scores: 1.48e-120 Length: 274
Pred. No.: 1072.50 Matches: 204
Score:

```

```

Percent Similarity: 85.56% Conservative: 27
Best Local Similarity: 75.56% Mismatches: 38
Query Match: 58.10% Indels: 1
DB: 3 Gaps: 1
US-10-698-228-2 (1-1053) x US-09-299-268-19 (1-274)
QY 94 GAGCCACTCTTAAGAAAGAGTTCTCGCCGGTTTGTCTATCTTTCCAAATCCAGTACCCCTGAT 153
Db 2 GluProIleLeuGlnGluSerAspSerArgPheValIlePheProIleLysTyrHisAsp 21
QY 154 ATTGGAATAATGTATAACAGGCACAGGCTTCCTTCTGGAACAGAGAGGTTCGACTTA 213
Db 22 IleTrpLysMetTyrLysGlnSerValAlaSerPheTrpThrValGluGluValAspLeu 41
QY 214 TCAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAAGTACTTCTCTCAC 273
Db 42 SerLysAspLeuAspAspTrpAspLysLeuThrLysAspGluLysTyrPheIleLysHis 61
QY 274 ATCTTAGCCCTTTTTCAGCCAGTGTGAATTTGTAATGAAATTTGGTGGAGCGCTTT 333
Db 62 IleuAlaPhePheAlaSerSerAspGlyIleValAsnGluAsnLeuAlaGluArgPhe 81
QY 334 AGTCAGGAGTGCAGGTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAATTTCTCATCGAG 393
Db 82 TyrValAspValGlnCysSerGluAlaArgCysPheTyrGlyPheGlnIleAlaMetGlu 101
QY 394 AATGTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAGATCCCAAGAAA 453
Db 102 AsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrValArgAspAsnIleGlu 121
QY 454 AGGGAATTTTATTGAATGCAATTGAAACCATGCCCTATGTTAAGAAAAAGCAGATTGG 513
Db 122 LysMetHisLeuPheAsnAlaIleGluThrMetGluCysValLysLysLysAlaAspTrp 141
QY 514 GCCTTGCATGTAGACAGATAGAAAATCTACTTTTGGGAAAGAGTGGTGGCCTTTCCT 573
Db 142 AlaArgLysTrpIleSerSerAsnLys--ValTyrGlyGluArgValValAlaPheAla 160
QY 574 GCTGTAGAGGAGTTTCTTCTCAGGATCTTTTGTCTGCTATATCTGCTGCTAAAGAGAGA 633
Db 161 AlaValGluGlyIlePhePheSerGlySerPheAlaAlaIlePheTrpIleLysLysArg 180
QY 634 GGTCTTATGCCAGGACTCACTTTTCCAATGAATCATCAGCAGATGAAGGACTTCAC 693
Db 181 GlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHis 200
QY 694 TGTGACTTTTCTGCTGCTGATGTTCCAAATCTAGTAAATAAGCCTTCAGAAAGAGGTC 753
Db 201 CysAspPheAlaCysLeuMetPheLysHisLeuHisProProSerLysGluValIle 220
QY 754 AGGGAGATCATGTTGATGCTGCAAAATTTGAGCAGGAGTTTAAACAGAGCCCTGGCA 813
Db 221 ThrSerIleIleIleAspAlaValAsnIleGluLysGluPheLeuThrValAlaIlePro 240
QY 814 GTTGGCCTCATTTGGAATGAAATTTGATGAAACAGTACATTTGAGTTGTAGCTGAC 873
Db 241 ValAspLeuIleGlyMetAsnCysCysLeuMetSerGlnIryIleGluPheValAlaAsp 260
QY 874 AGATTACTTGTGGAACCTTGGATTCTCAAG 903
Db 261 ArgLeuLeuThrGluLeuGlyCysGluLys 270

```

RESULT 8

```

US-08-136-743B-2
; Sequence 2, Application US/08136743B
; Patent No. 5459063
; GENERAL INFORMATION:
; APPLICANT: Barry S. Cooperman, Harvey Rubin,
; APPLICANT: Jerome Salem, and Alison L. Fisher
; TITLE OF INVENTION: "plasmodium falciparum Ribonu-
; TITLE OF INVENTION: cleotide Reductase, DNA Sequences Therefor and Peptide Inhibit
; TITLE OF INVENTION: Thereof"

```



```
QY 127 GTCATCTTTCCAAATCCAGTACCCTGATATTGGAAAAATGTATAAACAGGCACAGGCTTCC 186
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
112 ValMetPheProIleArgTyrHisGluIleTrpAsnPheTyrLysLysAlaGluAlaSer 131
QY 187 TTTCTGACAGCAGAGAGAGTGCAGTATCAAAAGGATCTCCCTCAGTGG---AACAGCTT 243
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
132 PheTrpThrAlaGluGluIleAspLeuSerLysAspLeuAspAspTrpAsnAsnLysLeu 151
QY 244 AAAGCAGATGAACTACTCTCTACATCTCTACATCTCTCTCTCTCTCTCTCTCTCTCT 303
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
152 AsnGluAsnGluArgTyrPheIleSerArgValLeuAlaPhePheAlaAlaSerAspGly 171
QY 304 ATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCTCGC 363
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
172 IleValGlyGluAsnLeuIleGluAsnPheSerThrGluValGlnLeuProGluAlaLys 191
QY 364 TGTTCATATGCTTTCAATTCATCGAGATGTTCACTCAGAGATGATACAGTATGCTG 423
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
192 SerPheTyrGlyPheGlnIleMetMetGluAsnIleHisSerGluThrTyrSerLeuLeu 211
QY 424 ATAGACACTTACATCAGAGATCCCAAGAAAGGAATTTTATTTAATCAATTTGAACC 483
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
212 IleGluThrTyrIleLysAspProGlnGluAlaAspTyrLeuPheAsnAlaIleAlaAsn 231
QY 484 ATGGCCTATGTTAAGAAAAAGCAGATTCGGCTTCGATGATGATGATGATGATGATGAT 543
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
232 IleProCysIleGlnLysLysAlaAspTrpAlaIleLysTrpIleGlnAspAspGluAla 251
QY 544 ACTTTTGGGGAAGAGTGTGGCTTTGCTGTAGAGAGGATTTTCTTCTCAGGATCT 603
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
252 LeuTyrGlyGluArgLeuValAlaPheAlaAlaValGluGlyIlePhePheSerGlySer 271
QY 604 TTTGCTGTATATTCTGGCTTAAGAGAGAGAGTCTTATGCCAGGACTCACTTTTCCAA 663
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
272 PheAlaSerIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSerAsn 291
QY 664 GAATCATCAGCAGAGATGAAGGACTTCACTGTGCTGTGCTGTGCTGTGCTGTGCTG 723
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
292 GluLeuIleCysArgAspGluGlyLeuHisThrAspPheAlaCysLeuLeuPheAlaHis 311
QY 724 TTAGTAATAAGCCTTCAGAGAAAGGGTCAGGGAGATCATTTGTTGATGCTGTCAAAAT 783
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
312 LeuGlnAsnArgProSerProGluIleValGluArgIleIleThrGluAlaValAspVal 331
QY 784 GAGCAGGAGTTTAAAGAGCCTTGGCCAGTTGGCTCATTTGGAATGAATGCAATTTT 843
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
332 GluLysGluTyrPheThrAspValLeuProValSerLeuLeuGlyLeuAsnCysAsnLeu 351
QY 844 ATGAAACAGTACATTCAGTTTCTA 867
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
352 MetCysProTyrIleGluPheLeu 359
```

RESULT 10

```
US-08-905-223-358
; Sequence 358, Application US/08905223
; Patent No. 6222029
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste D.
; APPLICANT: Duclert, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS
; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Knobbe, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
```

RESULT 11

```
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/905,223
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Israel, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 358:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 147 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; TISSUE TYPE: Brain
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: -101...-1
; IDENTIFICATION METHOD: Von Heijne matrix
; OTHER INFORMATION: score 4.4
; OTHER INFORMATION: seq ISHILAFFAASDG/IV
; US-08-905-223-358

Alignment Scores:
Pred. No.: 7,19e-84 Length: 147
Score: 768.00 Matches: 147
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 41.60% Indels: 0
DB: 3 Gaps: 0

US-10-698-228-2 (1-1053) x US-08-905-223-358 (1-147)
QY 1 ATGGCGACCCGAAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
DB 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACACGAAAGTGAATAAAGTCAATCAAGAGCCACTCTTAAGAAAGAGTCTCGC 120
DB 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSer 40
QY 121 CGGTTTGTCTATCTTCCAAATCCAGTACCCCTGATATTGGAAAAATGTATAAACAGC 180
DB 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAla 60
QY 181 GCTTCTCTTGGACAGCAGAGGCTGACTTATCAAAAGGATCTCCCTCACTGGAACAAG 240
DB 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsn 80
QY 241 CTTAAAGCAGATCAGAGTACTTCACTCTCAGATCTTAGCTTTTTCAGCCAGTCA 300
DB 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSer 100
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
DB 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGlu 120
QY 361 CGCTGTTTCTATCGCTTTCAAAATTTCTCATCAGAAATCTTCACTCAGAGATGTACAGTT 420
DB 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSer 140
QY 421 CTGATGACACTTACATCACA 441
DB 141 LeuIleAspThrTyrIleArg 147
```


US-09-270-767-46236
; Sequence 46236, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46236
; LENGTH: 149
; TYPE: PRT
; ORGANISM: *Drosophila melanogaster*
US-09-270-767-46236

Alignment Scores:
Pred. No.: 26-63 Length: 149
Score: 599.00 Matches: 114
Percent Similarity: 85.91% Conservative: 14
Best Local Similarity: 76.51% Mismatches: 21
Query Match: 32.45% Indels: 0
DB: 4 Gaps: 0

US-10-698-228-2 (1-1053) x US-09-270-767-46236 (1-149)
QY 586 GTTTCCTCTCAGGATCTTTGCTGCTATATCTGCTAAAGAGAGAGGTCTTATGCCA 645
Db 1 IIEPHEPHESErGlySerPheAlaSerIlePheTrpLeuLysLysArgGlyLeuMetPro 20
QY 646 GGACTCAGCTTTTCCAAATGAACTCATCAGCAGAGTGAAGGACTTCACCTGTGACTTTGCT 705
Db 21 GlyLeuThrPheSerAsnGluLeuSerArgGlyGlyLeuHisCysAspPheAla 40
QY 706 TGCCTGATGTTCCAAATCTAGTAATAAGCCTTCAGAAAGAGGCTCAGGAGATCATTT 765
Db 41 ValLeuMetPheGlnHisLeuValGlnArgProLysArgGluArgIleLeuGluIle 60
QY 766 GTTGATGCTGCAAAATGAGCAGAGGATTTTAAACGAAGCCCTGCGAGTGGCCCTCATTT 825
Db 61 ArgAspAlaValAlaIleGluGlnPheLeuThrAspAlaLeuProValAsnLeu 80
QY 826 GGAATGAATTCATTTTCATGAAACAGTACATTTGATGTTGATGCTCAGAGTACTTCTGTG 885
Db 81 GlyMetAsnCysAspLeuMetSerGlnTrpIleGluPheValAlaAspArgLeuLeuVal 100
QY 886 GAACCTGGATTCCTCAAGGCTTTTCAGGCAGAGAAATCCTTTTGAATTTATGAAACATTT 945
Db 101 GluLeuGlyValGlyLysIleTrpAsnThrLysAsnProPheAsnPheMetGluMetIle 120
QY 946 TCTTTAGAGGAAACAAATTTCTTTGAGAACGAGTTCAGAGTATCAGCGTTTGTGCA 1005
Db 121 SerLeuAspGlyLysThrAsnPhePheGluLysValGlyGluTrpGlnArgMetGly 140
QY 1006 GTTATGGCAGAAACACACAGATACCTC 1032
Db 141 ValValSerAsnProLeuAspAsnVal 149

RESULT 12
US-09-540-10006
; Sequence 10006, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: *Myxococcus xanthus* Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883

; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 10006
; LENGTH: 323
; TYPE: PRT
; ORGANISM: *Myxococcus xanthus*
US-09-540-10006

Alignment Scores:
Pred. No.: 1-15e-50 Length: 323
Score: 495.00 Matches: 113
Percent Similarity: 54.49% Conservative: 63
Best Local Similarity: 34.98% Mismatches: 125
Query Match: 26.81% Indels: 22
DB: 4 Gaps: 6

US-10-698-228-2 (1-1053) x US-09-540-10006 (1-323)

QY 145 TACCCTGATATTTGGAAAAATGTATAAACAGCAGCAGCGCTCTCTCTGGACAGCAGAAGAG 204
Db 3 TyrProGlnPhePheGluMetTyrArgAsnAlaIleLysAsnThrTrpThrValGluGlu 22
QY 205 GTCGACTTATCAAGGATCTCCCTCACTGG--AACAGCTTAAAGCAGATGAGAGTAC 261
Db 23 ValAspPheSerThrAspLeuValAspLeuArgSerLysMetThrAspAlaGluArgHis 42
QY 262 TTCACTCTCATCTTAGCTTTTTCACGCCAGTGAATTTGTAATGTAATAAATTTG 321
Db 43 LeuIleHisArgLeuValAlaPhePheAlaThrGlyAspSerIleValGlyAsnAsnLeu 62
QY 322 GTCGAGCGCTTTAGTCAGAGGTCAGGTTCCAGAGCGCTCGCTGTTCTATGGCTTTCAA 381
Db 63 ValLeuAsnLeuTyrLysHisLeuAsnAlaProGluAlaArgMetTyrLysSerArgGln 82
QY 382 ATTCTCATCGAGAATGTTCTACTCAGAGATGTACAGTTTGTGATAGACACACTTACATCAGA 441
Db 83 LeuTyrGluGluAlaLeuHisValGlnPheTyrLeuThrLeuLeuAspThrTyrValPro 102
QY 442 GATCCCAAGAAAAGGAAATTTTATTAAATGCAATTTGAAACCATGCGCTATGTTAAGAAA 501
Db 103 AspProAlaGluArgAlaMetAlaPheAlaAlaValAspAsnIleProSerIleGlnArg 122
QY 502 AAAGCAGATTTGGCGCTTGCATGCGATA-----GCAGATAGAAAATCTACT 546
Db 123 LysAlaGlnPheCysMetLysTrpMetAspSerIleGlnGlyLeuAspThrLeuGlnThr 142
QY 547 TTTGGGGAAGA-----GTGCTGCGCTTTTGCT--GCTGTAGAAGA 585
Db 143 LysAlaHisArgArgGlnPheLeuLeuAsnLeuIleCysPheAlaGlyCysIleGluGly 162
QY 586 GTTTTCTTCTCAGGATCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCA 645
Db 163 LeuPhePhePheAlaAlaPheAlaTyrValTyrPheLeuArgSerLysGlyLeuLeuAsn 182
QY 646 GGACTCAGCTTTTCCAAATGAACTCATCAGCAGAGATGAAGGACTTCAGCTGCTTCTGCT 705
Db 183 GlyLeuAlaAlaGlyThrAsnTrpValPheArgAspGluSerAlaHisMetAlaPheAla 202
QY 706 TGCCTGATGTTCCAATAC-----TTAGTAAATAAGCTTTCAGAA 744
Db 203 PheGluSerIleGlnValAlaAlaArgLysGluGluProAspLeuPheAspAlaGlnMetGlu 222
QY 745 GAAAGGCTCAGGAGATCATTTGATGCTGTCAAAATTTGACGAGGAGTTTTTTAACAGAA 804
Db 223 ArgAspValValAlaMetLeuArgGluAlaValGluCysGluThrGlnPheAlaGlnAsp 242
QY 805 GCCTTGCCAGTTCGCTCATTTGAAATGCAATTTGATGAAACACAGTACATGAGTTT 864
Db 243 LeuLeuSerGlyGlyValMetGlyLeuSerValGlnAspMetArgGlyTyrLeuGluTrp 262
QY 865 GTAGCTGACAGATTTACTTGTGGAACCTTGGATTTCTCAAGGTTTTCAGGCGAGAAATCCT 924
Db 263 ValAlaAspGlnArgLeuGlnMetLeuGlyMetThrProValPheGlyThrLysAsnPro 282

QY 925 TTTGATTTTATGGAACCATCTTTCTAGAGGAAAAACAAATTTCTTTGAGAAACGAGTT 984
Db 283 LeuHisPheMetAspLeuGlnAspValGlnGluLeuThrAsnPhePheGluArgVal 302
QY 985 TCAGAGTATCAGCGTTTTCAGTATGCGCAGAAACCCACAGATAAGCTCTTCACCTTGGAT 1044
Db 303 SerSerTyrGlnValAlaValGlyValGlyAlaAlaThrAspValVal-----LeuAsp 320
QY 1045 GCAGATTTT 1053
Db 321 AlaAlaPhe 323

RESULT 13
US-09-328-352-4714
; Sequence 4714, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; PRIORITY FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4714
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4714

Alignment Scores:
Pred. No.: 7,35e-37 Length: 445
Score: 381.50 Matches: 97
Percent Similarity: 49.57% Conservative: 76
Best Local Similarity: 27.79% Mismatches: 143
Query Match: 20.67% Indels: 33
DB: 4 Gaps: 10

US-10-698-228-2 (1-1053) x US-09-328-352-4714 (1-445)

QY 25 GCGGCGCGGTGGATCAGGATCAGAGATCATCTTCAGACACCAACCAAGTGAATAAAG 84
Db 98 AlaProGlyLeuGluGluLeuGluMetGlyAla-----GlnArgValGln 112
QY 85 TCAATGAAGACCCATCTCTAAGAAGAGTTCTCCCGGTTTGTCTATCTTTCCAAATCCAG 144
Db 113 ValAspAspLysAlaMetileAsnCysArgAlaAspLeuAsnGlnLeuValProPheLys 132
QY 145 TACCTGTATTTGAAATATGTAATAACAGGCACAGGCTTCCTCTGGACAGCAGAGAG 204
Db 133 TyrGluTrpAlaTrpGlnLysTyrLeuAspGlyCysAlaAsnHisTrpMetProGlnGlu 152
QY 205 GTCGATTTATCAAGATCTCCCTCACTGG-----AACAGCTTAAAGCAGATGAG 255
Db 153 ValAsnMetAsnHisAspIleAlaLeuTrpLysSerGluAsnGlyLeuThrGluAspGlu 172
QY 256 AAGTACTTCTCATCTCACAATCTTAGCTTTTTCAGCCAGGATGGAATTTGTAATGAA 315
Db 173 ArgThrIleValMetArgSerLeuGlyPhePheSerThrAlaAspSerLeuValAlaAsn 192
QY 316 AATTGGTGGAGCGCTTAGTACGAGGTGCGAGTTCCAGAGCTCGCTGTTTCTATGGC 375
Db 193 AsnLeuValLeuAlaIleTyrArgHisIleThrAsnProGluCysArgGlnTyrIleLeu 212
QY 376 TTTCAAATCTCATCGAATCTTCACACAGAGATGTACAGTTTGTGATAGACATTAC 435
Db 213 ArgGlnAlaPheGluGluAlaIleHisThrHisAlaTyrGlnTyrCysIleGluSerLeu 232
QY 436 ATCAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTTGAAACCAATGCGCTATGTT 495
Db 233 GlyMetAspGluGlyGlu-----ValPheAsnMetTyrArgGluIleProSerVal 249

QY 496 AAGAAAAAAGCAGATTGGCGCTTGCATGG-----ATAGCAGATAGAAAAATCTACT 546
Db 250 AlaArgLeuAlaAlaTrpGlyLeuLysTyrThrGlnSerLeuSerAspProThrPheHis 269
QY 547 TTTGGG-----GAAAGAGTGTGGCTTTTGGCTGTGTA 579
Db 270 ThrGlyThrProGluAsnAspGlnArgLeuLeuArgAsnLeuIleAlaPheTyrCysVal 289
QY 580 ---GAAGGAGTTTCTCTCAGGATCTTTTGTGCTATATCTGCTTAAAGAAAGAGGT 636
Db 290 LeuGluGlyIlePhePheTyrCysGlyThrGlnIleLeuSerMetGlyArgAsn 309
QY 637 CTTATGCCAGGACTCATTCTTTTCCAATGAACACTCATCAGCAGATGAAGGACTTCACTGT 696
Db 310 LysMetAsnGlyValAlaGluGlnPheGlnTyrIleLeuArgAspGluSerMetHisLeu 329
QY 697 GACTTTGTCTGCTGATGTTTC---CAATCTTGTAGTAATAAGCTTCA----- 741
Db 330 AsnPheGlyIleAspMetIleAsnGlnIleLysIleGluAsnProHisLeuTrpThrAla 349
QY 742 -----GAAGAAAGGCTCAGGAGATCATTTGTGATCTCTCAAAATGACGAGGTTT 795
Db 350 GluPheGlnGlnGluValIleGlnMetIleLeuGluGlyThrMetLeuGluIleGluTyr 369
QY 796 TTAACAGAAAGCCTTGGCAGTTGGCCTCATTTGGAATGAATTCATTTTATGATGAAACAGTAC 855
Db 370 AlaArgAspThrMetProArgGlyValLeuGlyMetAsnAlaSerMetMetGluGluTyr 389
QY 856 ATTGAGTTTGTAGCTGCAGATTTACTTGTGGAATCTTCAAGATTTTCAAGTTTTCAG--- 912
Db 390 LeuLysPheIleCysAsnArgArgLeuSerGlnLeuGlyLeuProGluGlnPheAlaGly 409
QY 913 GCAGAAAACTCTTTTGTATTTATG---GAAACATTTCTTTAGAGGAAAAACAATTTTC 969
Db 410 ValThrAsnProPheAlaTrpMetSerGluMetMetAspLeuArgLysGluLysAsnPhe 429
QY 970 TTTGAGAAACGAGTTTTCAGAGTATCAG 996
Db 430 PheGluThrArgValThrAspTyrGln 438

RESULT 14
US-09-252-991A-19358
; Sequence 19358, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIORITY FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19358
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19358

Alignment Scores:
Pred. No.: 2.32e-36 Length: 476
Score: 377.50 Matches: 100
Percent Similarity: 45.87% Conservative: 72
Best Local Similarity: 26.67% Mismatches: 154
Query Match: 20.45% Indels: 49
DB: 4 Gaps: 11

US-10-698-228-2 (1-1053) x US-09-252-991A-19358 (1-476)

QY 4 GCGCACCAGAAAGCGGAGCGCGGCTGGATCAGGATGAG-----AGATCA 54


```
Db      243  LeuArgAspGluThrIleHisLeuAsnPheGlyIleAspLeuIleAsnGlyIleLysGlu 262
QY      733  AAGCCTTCAGAA-----GAAAGGTCAGGAGATCATTTGTTGAT 771
      :::: ||| :::: :::: :::: :::: ::::
Db      263  GluAsnProGluValTrpThrThrGluLeuGlnGluIleValAlaLeuIleGluLys 282
QY      772  GCTGTCAAATTCAGCAGAGTTTTTAACAGAAAGCCTTGCCAGTTGGCCTCATTTGGAATG 831
      ||||| :::: :::: :::: :::: :::: ::::
Db      283  AlaValGluLeuGluIleGluTyrAlaLysAspCysLeuProArgGlyIleLeuGlyLeu 302
      :::: :::: :::: :::: :::: ::::
QY      832  AATTCATTTTCATCAACACAGTACATTGAGTTGTAGCTACAGATTACTTGTGGAACCT 891
      :::: ||| :::: :::: :::: :::: ::::
Db      303  ArgSerSerMetPheIleAspTyrValArgHisIleAlaAspArgLeuGluArgIle 322
QY      892  GGATTCCTCAAGGTTTTTCAGGCAGAAAATCCTTTTGATTTTATG--GAAAACATTTCT 948
      ||| :::: :::: :::: :::: :::: ::::
Db      323  GlyLeuLysProIleTyrHisSerArgAsnProPheProTrpMetSerGluThrMetAsp 342
QY      949  TTAGAAGGAAACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAG 996
      ||| :::: :::: :::: :::: :::: ::::
Db      343  LeuAsnLysGluLysAsnPheGluThrArgValThrGluTyrGln 358
```

Search completed: October 30, 2005, 06:36:12
Job time : 24.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 30, 2005, 06:34:06 ; Search time 65.75 Seconds
(without alignments)
13390.093 Million cell updates/sec

Title: US-10-698-228-2

Perfect score: 1846

Sequence: 1 atgggagccgcgaagggc.....tcaccttgatgcagatttt 1053

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1865214 seqs, 418043040 residues

Total number of hits satisfying chosen parameters: 3730428

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlh
-Q=/cgn2_1/USPRO.spool/US10698228/runat_26102005_100610_4101/app_query.fasta_1.2446
-DB=Published Applications AA -QMTX=fastan -SUFFIX=n2p.rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFWT=pro -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10698228 @cgn1_113 @runat_26102005_100610_4101
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOC=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pcp.*
- 2: /cgn2_6/ptodata/1/pubpaa/ECT_NEW_PUB.pcp.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pcp.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pcp.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pcp.*
- 6: /cgn2_6/ptodata/1/pubpaa/ECTUS_PUBCOMB.pcp.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pcp.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pcp.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pcp.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pcp.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pcp.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pcp.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pcp.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pcp.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pcp.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pcp.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pcp.*
- 18: /cgn2_6/ptodata/1/pubpaa/US10F_NEW_PUB.pcp.*
- 19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pcp.*
- 20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pcp.*
- 21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pcp.*
- 22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1821	98.6	351	15	US-10-698-228-1 Sequence 1, Appli
2	1454	78.8	389	15	US-10-698-228-5 Sequence 5, Appli
3	1454	78.8	389	16	US-10-408-765A-559 Sequence 559, App
4	1454	78.8	389	16	US-10-733-878-457 Sequence 457, App
5	1454	78.8	389	18	US-10-756-149-4899 Sequence 4899, Ap
6	1454	78.8	413	9	US-09-925-301-1347 Sequence 1347, Ap
7	1454	78.8	453	18	US-10-220-335-686 Sequence 686, App
8	1454	78.8	453	18	US-10-450-763-45646 Sequence 45646, A
9	1425.5	77.2	386	15	US-10-403-571-76 Sequence 76, Appl
10	1348	73.0	393	20	US-11-097-143-19182 Sequence 19182, A
11	1242	67.3	430	15	US-10-369-493-4122 Sequence 4122, Ap
12	1230	66.6	405	14	US-10-128-714-8245 Sequence 8245, Ap
13	1214	65.8	381	15	US-10-369-493-5686 Sequence 5686, Ap
14	1192.5	64.6	381	14	US-10-128-714-3245 Sequence 3245, Ap
15	1192	64.6	413	14	US-10-032-585-7537 Sequence 7537, Ap
16	1176.5	63.7	321	15	US-10-369-493-10365 Sequence 10365, A
17	1169.5	63.4	399	15	US-10-369-493-22170 Sequence 22170, A
18	1161.5	62.9	331	15	US-10-369-493-22565 Sequence 22565, A
19	1161.5	62.9	401	15	US-10-320-797-3145 Sequence 3145, Ap
20	1147.5	62.2	329	15	US-10-424-599-215866 Sequence 215866, Sequence 215845,
21	1133.5	61.4	717	16	US-10-425-115-215845 Sequence 215845,
22	1132.5	61.3	345	16	US-10-437-963-113592 Sequence 113592, A
23	1128.5	61.1	339	16	US-10-767-701-45287 Sequence 45287, A
24	1125.5	61.0	339	16	US-10-425-115-247984 Sequence 247984, Sequence 269685,
25	1114.5	60.4	329	15	US-10-424-599-183313 Sequence 183313, Sequence 200125,
26	1097.5	59.5	339	15	US-10-424-599-183313 Sequence 200125,
27	1092.5	59.2	400	16	US-10-425-115-200125 Sequence 200125,
28	1061	57.5	347	16	US-10-739-930-6922 Sequence 6922, Ap
29	916	49.6	270	18	US-10-220-335-342 Sequence 342, App
30	871.5	47.2	344	15	US-10-369-493-13401 Sequence 13401, A
31	840	45.5	254	16	US-10-425-115-215843 Sequence 215843, A
32	805	43.6	345	15	US-10-369-493-215847 Sequence 215847, A
33	784	42.5	225	16	US-10-425-115-215840 Sequence 215840, Sequence 2195,
34	701	38.0	138	15	US-10-264-237-2195 Sequence 2195, Ap
35	680	36.8	335	18	US-10-450-763-45645 Sequence 45645, A
36	662	35.9	181	18	US-10-450-763-37219 Sequence 37219, A
37	564.5	30.6	338	14	US-10-156-761-10564 Sequence 10564, A
38	544	29.5	156	15	US-10-424-599-215870 Sequence 215870, A
39	491	26.6	307	15	US-10-369-493-19363 Sequence 19363, A
40	456.5	24.7	145	16	US-10-767-701-45286 Sequence 45286, A
41	397	21.5	398	15	US-10-282-122A-51357 Sequence 51357, A
42	384.5	20.8	415	15	US-10-282-122A-66252 Sequence 66252, A
43	384	20.8	403	15	US-10-282-122A-50118 Sequence 50118, A
44	383.5	20.8	415	15	US-10-282-122A-69814 Sequence 69814, A
45	380.5	20.6	416	15	US-10-282-122A-68043 Sequence 68043, A

ALIGNMENTS

RESULT 1

US-10-698-228-1
; Sequence 1, Application US/10698228
; Publication No. US2004007253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698, 228
; PRIOR FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 1

```
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-1

Alignment Scores:
Pred. No.: 1,7e-187 Length: 351
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.65% Indels: 0
DB: 15 Gaps: 0

US-10-698-228-2 (1-1053) x US-10-698-228-1 (1-351)

QY 1 ATGGCGACCCGGAAGCGCGGCTGGATCAGATGAGATCATCTTCA 60
DB 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAAGCAAGTGAATGAATCAATGAAGACCACTCCTAAGAAAGAGTTCTCGC 120
DB 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
QY 121 CGGTTTGTCTATCTTCCATCCAGTACCTGATATTTGGAAATGTATAAAGAGGACAG 180
DB 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCCTCTGACAGCAGAGAGGTGCGACTTATCAAAGGATCTCCCTCACTGGAACAAG 240
DB 61 AlaSerPheTrpAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAGCAGATGAGAGTACTTCTCATCTCATCTTACGCTTTTTCGACGCGATGAT 300
DB 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY 301 GCAATTTGTAAGTAAATTTGTGGAGCGCTTTAGTCAGGAGGTGCGAGGCT 360
DB 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATGGCTTTTCAAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 420
DB 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACATTATCATCAGAGATCCCAAGAAAGGAATTTTATTAATGCAATGAA 480
DB 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATTCGGCTTCGGATGGATAGCAGATAGAAA 540
DB 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAAGAGGAGTTTCTTCTCAGGA 600
DB 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGTCTTATGCGAGACTCAGCTTTTTC 660
DB 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACATCATCAGCAGATGAGAGCTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
DB 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAAATAAGCTTCAGAGAAAGGTCAGGAGATCATCTTGTGATGCTGTCAAA 780
DB 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
QY 781 ATTGAGCAGGAGTTTAAACAGAGGCTTGCAGAGTTGGCTCATTTGGAATGAATTCATT 840
DB 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
```

```
QY 841 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
DB 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTCAGGCAGAAAATCCTTTTGTGATTTTATGAAAAACATTTCTTTAGAAAGAAA 960
DB 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAAAATTTCTTTCAGAAACAGTTTCAGAGTATTCAGCGTTTTCAGTTATGCGCAGAAACC 1020
DB 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGCTTCTTCCCTTTGGATGAGATTTT 1053
DB 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 2
US-10-698-228-5
; Sequence 5, Application US/10698228
; Publication No. US20040072253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WO0P
; CURRENT APPLICATION NUMBER: US/10/698,228
; PRIOR FILING DATE: 2003-10-30
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 5
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-5

Alignment Scores:
Pred. No.: 9.03e-148 Length: 389
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: 15 Gaps: 0

US-10-698-228-2 (1-1053) x US-10-698-228-5 (1-389)

QY 91 GAAGAGCCACTCTCTAAGAAAGAGTCTTCGCCGGTTTGTTCATCTTTCCAATCCAGTACCCT 150
DB 69 AspGluProLeuLeuArgGluAsnProArgArgPheValIlePheProIleGluTyrHis 88
QY 151 GATATTGGAAAATGTATAAACAGGCACAGGCTTCTTCTGACAGAGAGAGGTTCGAC 210
DB 89 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpAlaGluGluValAsp 108
QY 211 TTATCAAGGATCTCCTCTGACCTGGAACAGCTTAAACAGATGAGAGTACTTCTCTCT 270
DB 109 LeuSerLysAspIleGlnHisTrpGluSerLysProGluGluArgTyrPheIleSer 128
QY 271 CACATCTTAGCCTTTTTCAGCAGCAGTGTGAATTCGAAATGAAATTTTGGTGAGCGC 330
DB 129 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148
QY 331 TTTTAGTCAGGAGGTGCGAGGTTCCAGAGGCTCGCTGTGTTTCTATGGCTTTCAAATTCATC 390
DB 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
```

```
QY 391 GAGATGTTCTCAGAGATGACAGTTTGTGATAGACACCTTACATCAGAGATCCCAAG 450
|||...|||
Db 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
|||...|||
QY 451 AAAAGGGAATTTTATTAATGCAATTAAGAACCATGCCCTATGTTTAAAGAAAAGCAGAT 510
|||...|||
Db 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 208
|||...|||
QY 511 TGGGCTCTCCGATGAGATAGACAGATGAAATCTACTTTTGGGGAAGAGTGGTGGCTTT 570
|||...|||
Db 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaIlePhe 228
|||...|||
QY 571 GTCTGCTAGAGAGATTTCTTCTCAGAGATCTTTTGTGCTATATCTCGCTATAATCTGGCTAAAGAAAG 630
|||...|||
Db 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 248
|||...|||
QY 631 AGAGGCTTATGCGCAGGACTCATTCTTTTCAATGAACTCAGCAGAGATGAGAGACTT 690
|||...|||
Db 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
|||...|||
QY 691 CACTGTGACTTTGCTTGGCTGATGTTTCCAAATAGTAAATAGAGCTTCCAGAGAAAGG 750
|||...|||
Db 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
|||...|||
QY 751 GTCCAGGAGATCATTTGCTGCTGATGTTTCCAAATAGTAAATAGAGCTTCCAGAGAAAGG 810
|||...|||
Db 289 ValArgGluIleIleLeuAsnAlaValArgIleGluGluPheLeuThrGluAlaLeu 308
|||...|||
QY 811 CCAGTTGGCCTCATTTGCAATGAAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 870
|||...|||
Db 309 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 328
|||...|||
QY 871 GACAGATTTACTTGTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 930
|||...|||
Db 329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348
|||...|||
QY 931 TTTATGGAACAACTTTCTTTAGAGGAAACAACTTTCTTTAGAGGAAACAACTTTCTTTAGAGG 990
|||...|||
Db 349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368
|||...|||
QY 991 TATCAGCGTTTTCAGTATAGGAGAAACCAACAGATTAACGCTTTCACCTTGGATGCAGAT 1050
|||...|||
Db 369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
|||...|||
```

RESULT 3

```
US-10-408-765A-559
; Sequence 559, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 559
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-559
```

Alignment Scores:

```
Pred. No.: 9.03e-148 Length: 389
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: 16 Gaps: 0
```

US-10-698-228-2 (1-1053) x US-10-408-765A-559 (1-389)

```
QY 91 GAGAGCCACTCTCTAAGAAAGAGTTCTCGCGGTTTGTCTATCTTCCAAATCCAGTACCCT 150
|||...|||
Db 69 AspGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGluTyrHis 88
|||...|||
QY 151 GATATTTGGAAATGATATAACAGGCACAGCTTCTCTGACAGCAGAGAGAGTCCAC 210
|||...|||
Db 89 AspIleTrpGlnMetTyrLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 108
|||...|||
QY 211 TTATCAAGAGATCTCCCTCACTGTGAACAAGCTTAAAGCAGATGAGAGTACTTCTATCTCT 270
|||...|||
Db 109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128
|||...|||
QY 271 CACATCTTACCTTTTTCAGCCAGTGTGAAATGTAATGAAATTTGGTGGAGCGC 330
|||...|||
Db 129 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148
|||...|||
QY 331 TTTAGTCAGAGGTGCGAGTTCAGAGCTCGCTGTTTCTATGCTTTCAAATTTCTCATC 390
|||...|||
Db 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
|||...|||
QY 391 GAGAATGTTTCACTCAGAGATGTACAGTTGTCTGTATAGACACTTACATCAGAGATCCCAAG 450
|||...|||
Db 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
|||...|||
QY 451 AAAAGGGAATTTTATTAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 510
|||...|||
Db 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 208
|||...|||
QY 511 TGGGCTCTCCGATGAGATAGACAGATGAAATCTACTTTTGGGGAAGAGTGGTGGCTTT 570
|||...|||
Db 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaIlePhe 228
|||...|||
QY 571 GCTGCTGTAAGAGAGTTTCTTCTCAGAGATCTTTTGTGCTATATCTGCTATAATCTGGCTAAAGAG 630
|||...|||
Db 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 248
|||...|||
QY 631 AGAGGCTTATGCGCAGGACTCATTCTTTTCCAAATGCAATCAGCAGAGATGAGAGACTT 690
|||...|||
Db 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
|||...|||
QY 691 CACTGTGACTTTGCTTGGCTGATGTTTCCAAATAGTAAATAGAGCTTCCAGAGAAAGG 750
|||...|||
Db 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
|||...|||
QY 751 GTCAGGAGATCATTTGCTGCTGATGTTTCCAAATGCAATGCAATGCAATGCAATGCAAT 810
|||...|||
Db 289 ValArgGluIleIleLeuAsnAlaValArgIleGluGluPheLeuThrGluAlaLeu 308
|||...|||
QY 811 CCAGTTGGCCTCATTTGCAATGAAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 870
|||...|||
Db 309 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 328
|||...|||
QY 871 GACAGATTTACTTGTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 930
|||...|||
Db 329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348
|||...|||
QY 931 TTTATGGAACAACTTTCTTTAGAGGAAACAACTTTCTTTAGAGGAAACAACTTTCTTTAGAGG 990
|||...|||
Db 349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368
|||...|||
QY 991 TATCAGCGTTTTCAGTATAGGAGAAACCAACAGATTAACGCTTTCACCTTGGATGCAGAT 1050
|||...|||
Db 369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
|||...|||
```

```
QY 1051 TTT 1053
Db 389 Phe 389

RESULT 4
US-10-733-878-457
; Sequence 457, Application US/10733878
; Publication No. US20040224408A1
; GENERAL INFORMATION:
; APPLICANT: Jean-Philippe Girard
; APPLICANT: Francois Amalric
; APPLICANT: Myriam Roussigne
; APPLICANT: Thomas Clouaite
; TITLE OF INVENTION: THAP PROTEINS AS NUCLEAR RECEPTORS FOR
; TITLE OF INVENTION: CHEMOKINES AND ROLES IN TRANSCRIPTIONAL REGULATION, CELL
; TITLE OF INVENTION: CHEMOKINES AND ROLES IN TRANSCRIPTIONAL REGULATION, CELL
; FILE REFERENCE: BIOBANK.012A
; CURRENT APPLICATION NUMBER: US/10/733,878
; PRIOR FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 60/432699
; PRIOR FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/485027
; PRIOR FILING DATE: 2003-07-03
; NUMBER OF SEQ ID NOS: 535
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 457
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-733-878-457

Alignment Scores:
Pred. No.: 9.03e-148 Length: 389
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: 16 Gaps: 0

US-10-698-228-2 (1-1053) x US-10-733-878-457 (1-389)

QY 91 GAAGGCCACTCTTAAGAAGAGTTCTCGCGGTTGTGTCATCTTTCCAATCCAGTACCCT 150
Db 69 AspGluProLeuLeuArgGluAsnProArgArgPheValIlePheProIleGluTyrHis 88

QY 151 GATATTGGAAATGTATAACAGGCACAGCGTCTCTTCGACAGCAGAGAGGTGCGAC 210
Db 89 AspIleTyrGlnMetTyrLysLysAlaGluAlaSerPheTyrThrAlaGluGluValAsp 108

QY 211 TTATCAAGGATCTCCCTCAGTGAACAGCTTAAGCAGATGAGAGTACTTCTCTCT 270
Db 109 LeuSerLysAspIleGlnHisTyrGluSerLeuLysProGluGluArgTyrPheIleSer 128

QY 271 CACATCTTAGCCTTTTTCAGCCAGTGAATGGAATGTAATGAAATTTGGTGGAGCGC 330
Db 129 HisValLeuAlaPhePheAlaIleAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148

QY 331 TTATGTCAGGAGTGCAGGTTCCAGAGGCTCGCTGTTCTATGGCTTTCAAAATTTCTATC 390
Db 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168

QY 391 GAGATGTTCACTCAGAGATGATACAGTTGCTGATAGACACTTACATCAGAGATCCCAAG 450
Db 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188

QY 451 AAAAGGGAATTTTATTTAATCAATGAAACCATGCCCTATCTTTAAGAAAAAGCAGAT 510
Db 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 208

QY 511 TGGGCTTCGATGAGATAGACATAGAAAATCTACTTTTGGGAAAGAGTGGTGGCCTTT 570
Db 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValValAlaPhe 228

571 GCTGCTGTAGAAGGAGTTTCTTCTCAGGATCTTTTCTGCTATATTTCTGCTAAAGAAG 630
Db 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 248

QY 631 AGAGTCTTATGCCAGGACTCACTTTTTCAAATCATCATCAGCAGAGATGAAGGACTT 690
Db 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268

QY 691 CACTGTGACTTTGCTTCCCTGATGTTCCCAATACTAGTAATAAGCCTTCAGAGAAGG 750
Db 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288

QY 751 GTCAGGAGATCATTTGATGCTGCAAAATTTGAGCAGGAGTTTAAACAGAGCCCTTG 810
Db 289 ValArgGluIleIleIleAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 308

QY 811 CCAGTTGGCTCATTCGAATGCAATTTGATGAAACAGTACATGATGTTGTAGCT 870
Db 309 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 328

QY 871 GACAGATTACTTGTGAACCTTGAATTTCAAGGTTTTCAGGACGAGAAATCCTTTGAT 930
Db 329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348

QY 931 TTTATCGAAAAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACGAGTTTCAGAG 990
Db 349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368

QY 991 TATCAGCGTTTTCAGTTATGCGAGAAACCAAGATAACCTCTTCACCTTGGATGCGAGAT 1050
Db 369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388

QY 1051 TTT 1053
Db 389 Phe 389

RESULT 5
US-10-756-149-4899
; Sequence 4899, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
; TITLE OF INVENTION: METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/756,149
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4899
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-756-149-4899

Alignment Scores:
Pred. No.: 9.03e-148 Length: 389
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: 18 Gaps: 0

US-10-698-228-2 (1-1053) x US-10-756-149-4899 (1-389)

QY 91 GAAGGCCACTCTTAAGAAGAGTTCTCGCGGTTGTGTCATCTTTCCAATCCAGTACCCT 150
Db 69 AspGluProLeuLeuArgGluAsnProArgArgPheValIlePheProIleGluTyrHis 88

QY 151 GATATTGGAAATGTATAACAGGCACAGCGTCTCTTCGACAGCAGAGAGGTGCGAC 210
Db 89 AspIleTyrGlnMetTyrLysLysAlaGluAlaSerPheTyrThrAlaGluGluValAsp 108

QY 211 TTATCAAGGATCTCCCTCAGTGAACAGCTTAAGCAGATGAGAGTACTTCTCTCT 270
Db 109 LeuSerLysAspIleGlnHisTyrGluSerLeuLysProGluGluArgTyrPheIleSer 128

QY 271 CACATCTTAGCCTTTTTCAGCCAGTGAATGGAATGTAATGAAATTTGGTGGAGCGC 330
Db 129 HisValLeuAlaPhePheAlaIleAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148

QY 331 TTATGTCAGGAGTGCAGGTTCCAGAGGCTCGCTGTTCTATGGCTTTCAAAATTTCTATC 390
Db 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168

QY 391 GAGATGTTCACTCAGAGATGATACAGTTGCTGATAGACACTTACATCAGAGATCCCAAG 450
Db 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188

QY 451 AAAAGGGAATTTTATTTAATCAATGAAACCATGCCCTATCTTTAAGAAAAAGCAGAT 510
Db 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 208

QY 511 TGGGCTTCGATGAGATAGACATAGAAAATCTACTTTTGGGAAAGAGTGGTGGCCTTT 570
Db 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValValAlaPhe 228
```



```
Db      89 AsplleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 108
QY      211 TTATCAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAGTACTTTCATCTCT 270
Db      109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128
QY      271 CACATCTTAGCCCTTTTTCAGCCAGTGTGAATTTGAAATTTGAAATTTGTTGGAGCGC 330
Db      129 HisValLeuAlaPhePheAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148
QY      331 TTATGTCAGGAGTGCAGGCTTCAGAGGCTCGCTGTTTTCATGGCTTTCAAATTTCTCATC 390
Db      149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
QY      391 GAGATGTTCACTCAGAGATGTACAGTGTGCTGATAGACACTTACATCAGAGATCCCAAG 450
Db      169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
QY      451 AAAAGGGAATTTTATTAATCAATGAAACCATGCCCTATGTTTAAAGAAAAAGCAGAT 510
Db      189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 208
QY      511 TGGGCTTCGCGATGATAGATAGAGAAATCTACTTTTGGGAAAGAGTGTGGCTTT 570
Db      209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaPhe 228
QY      571 GCTGCTGTAGAGGATTTTCTTCTCAGGATCTTTGCTGCTGCTATATTCGCTTAAAGAG 630
Db      229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLys 248
QY      631 AGAGGTCTTATGCCAGGACTCACTTTTCCAAATGAACTCATCAGCAGAGATGAAGACTT 690
Db      249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
QY      691 CACTGTGACTTTGCTTGCCTGATGTTCCAAATCTTAGTAAATAAGCCTTCAGAGAAAGG 750
Db      269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisIleProSerGluGluArg 288
QY      751 GTCAGGAGATCATTTGCTGATGCTGCAAAATTTGAGCAGAGTGTTTTAAACAGAGCTTG 810
Db      289 ValArgGluIleIleAlaValArgIleGluGluGluPheLeuThrGluAlaLeu 308
QY      811 CCAGTTGGCTCATTTGGAATGAATTCATTTGATGAAACAGTACATTCAGTTGTAGCT 870
Db      309 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 328
QY      871 GACAGATTACTTGTGAACTTGGAATCTCAAGGTTTTCAGGCAAAAATCTTTTGAT 930
Db      329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348
QY      931 TTTATGGAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACGAGTTTCAGAG 990
Db      349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368
QY      991 TATCAGCTTTTGCAGTTATGTCAGAAACCAAGATTAAGTCTTCACTTGCATGAGAT 1050
Db      369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
QY      1051 TTT 1053
Db      389 Phe 389
```

RESULT 6

```
US-09-925-301-1347
; Sequence 1347, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
```

```
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1347
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1347

Alignment Scores:
Pred. No.:      9,23e-148      Length:      413
Score:          1454.00      Matches:    269
Percent Similarity: 94.08%      Conservative: 33
Best Local Similarity: 83.80%      Mismatches: 19
Query Match:     78.76%      Indels:     0
DB:              0          Gaps:          0
```

US-10-698-228-2 (1-1053) x US-09-925-301-1347 (1-413)

```
QY      91 GAAGGCCACTCTCTAAGAAAGAGTTCTGCCGGTTTGTCTCATCTTTCATCCAGTACCTC 150
Db      93 AspGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGluTyrHis 112
QY      151 GATATTTGGAATGTATAACAGGACAGGCTTCTCTCGACAGCAGAGAGGTCCGAC 210
Db      113 AsplleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 132
QY      211 TTATCAAGGATCTCCCTCACTCGAACCAAGCTTAAAGCAGATGAGAGTACTTTCATCTCT 270
Db      133 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 152
QY      271 CACATCTTAGCCCTTTTTCAGCCAGTGTGAATTTGAAATTTGAAATTTGTTGGAGCGC 330
Db      153 HisValLeuAlaPhePheAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 172
QY      331 TTATGTCAGGAGTGCAGGTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAATTTCTCATC 390
Db      173 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 192
QY      391 GAGATGTTCACTCAGAGATGTACAGTTTCTGATGACACTTACATCAGAGATCCCAAG 450
Db      193 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 212
QY      451 AAAAGGGAATTTTATTAATGCAATTTGAAACCATGCCCTATGTTTAAAGAAAAAGCAGAT 510
Db      213 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 232
QY      511 TGGGCTTCGCGATGATAGATAGAAATCTACTTTTGGGAAAGAGTGTGGCTTTC 570
Db      233 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaPhe 252
QY      571 GCTGCTGTAGAGGATTTTCTTCTCAGGATCTTTTCTGCTGCTATATTCGCTTAAAGAG 630
Db      253 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLys 272
QY      631 AGAGGTCTTATGTCAGGACTCACTTTTTCAAATGAACTCATCAGCAGAGATGAAGGACTT 690
Db      273 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 292
QY      691 CACTGTGACTTTGCTTGCCTGATGTTCCAAATCTTAGTAAATAAGCCTTCAGAGAAAGG 750
Db      293 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 312
QY      751 GTCAGGAGATCATTTGCTGATGCTGCAAAATTTGAGCAGAGTGTTTTAAACAGAGCTTG 810
Db      313 ValArgGluIleIleAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 332
QY      811 CCAGTTGGCTCATTTGGAATGAATTCATTTGATGAAACAGTACATTCAGTTGTAGCT 870
Db      333 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 352
```

```
QY 871 GACAGATTACTTGTGGAAGTCTTCAAAAGTTTTCAGCGAGAAATCCTTTTGAT 930
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
353 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 372
QY 931 TTTATGGAACAACTTCTTAGAGGAAAAACAAATCTCTTTCAGAAACGAGTTTCAGAG 990
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
373 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 392
QY 991 TATCAGCGTTTTCAGTGTATGCGAGAAACACAGATAAAGCTCTTCACTTGGATGCAGAT 1050
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
393 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 412
QY 1051 TTT 1053
Db |||
413 Phe 413

RESULT 7
US-10-220-335-686
; Sequence 686, Application US/10220335
; Publication No. US20050175607A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-041
; CURRENT APPLICATION NUMBER: US/10/220,335
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 09/664,641
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/616,807
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 09/597,707
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/577,409
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: 09/515,126
; PRIOR FILING DATE: 2000-02-28
; NUMBER OF SEQ ID NOS: 688
; SOFTWARE: Custom
; SEQ ID NO 686
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-220-335-686

Alignment Scores:
Pred. No.: 9,56e-148 Length: 453
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: 18 Gaps: 0

US-10-698-228-2 (1-1053) x US-10-220-335-686 (1-453)
QY 91 GAAGAGCCACTCTTAAGAAGAGTCTCGCGGTTTGTCACTCTTCCAAATCCAGTACCCT 150
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
133 AppGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGluThrHis 152
QY 151 GATATTGGAATAATGATTAACAGGCACAGCGTCTCTTCGACAGCAGAAGAGGTCGAC 210
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
153 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 172
QY 211 TTATCAAAAGATCTCCCTACTCGAACACAGCTTAAGACGATGAGAAGTACTCTCTCT 270
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
173 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 192
QY 271 CACATCTTAGCCCTTTTTCAGCCAGTGTGAATGTAATCAAAATTTGGTGGAGCGC 330
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
193 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 212
QY 331 TTATAGTCAGGAGTGCAGGTTTCAGAGGCTCGCTGTTCTTATGGCTTCAAATTTCTCATC 390
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
213 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 232
```

```
QY 391 GAGAAATGTTCACTCAGAGATGTACAGTTGCTGATGACACTTTACATCAGAGATCCCAAG 450
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
233 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 252
QY 451 AAAAGGGAATTTTATTAATGCAATTGAACCAATGCGCTATGTTAAGAAAAAGCAGAT 510
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
253 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 272
QY 511 TGGGCTTTCGGATGGATAGCAGATAGAAATCTACITTTGGGGAAGAGTGGTGGCTTT 570
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
273 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaPhe 292
QY 571 GCTGCTGTAGAAGAGTCTTCTCTCAGGATCTTTTCTGCTCTATATTTCTGGCTAAAGAG 630
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
293 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 312
QY 631 AGAGTCTTTATGCCAGACTCCTTTTCCAAATCAATCAGCAGAGATGAAGGCTT 690
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
313 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 332
QY 691 CACTGTGACTTTGCTTGCCTGATGTTCCCAATCTAGTAATTAAGCCTTCAGAAAGAGG 750
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
333 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 352
QY 751 GTCAGGAGATCATTTGTTGATGCTGTCAAATTTGAGCAGGAGTCTTTTAAACAGAGCCTTG 810
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
353 ValArgGluIleIleIleAlaValArgIleGluGluPheLeuThrGluAlaLeu 372
QY 811 CCAGTTGGCTCTCATTTGGAATGAAATTTGATGAACACAGTACATTTGATGCT 870
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
373 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 392
QY 871 GACAGATTACTTGTGAACTTGTCAAGTCTCAAGTCTTTCAGCAGCAAAATCCTTTGAT 930
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
393 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 412
QY 931 TTTATGGAACAACTTCTTTAGAGGAAAAACAAATTTCTTCAGAAACGAGTTTCAGAG 990
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
413 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 432
QY 991 TATCAGCGTTTTCAGTGTATGCGAGAAACACAGATAAAGCTCTTCACTTGGATGCAGAT 1050
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
433 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 452
QY 1051 TTT 1053
Db |||
453 Phe 453

RESULT 8
US-10-450-763-45646
; Sequence 45646, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 45646
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
```

; LOCATION: (148)...(203)
; OTHER INFORMATION: Ribonucleotide reductase small subunit proteins domain
; OTHER INFORMATION: identified by eMATRIX, accession number BL00368A, p-value=1.000e-
; OTHER INFORMATION: 40, raw score of 36.98
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (167)...(440)
; OTHER INFORMATION: Ribonucleotide reductases domain identified by Pfam,
; OTHER INFORMATION: accession name ribonuc_red, E-value=9e-174, Pfam score of 507.8
US-10-450-763-45646

Alignment Scores:

Pred. No.: 9,56e-148 Length: 453
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: 18 Gaps: 0

US-10-698-228-2 (1-1053) x US-10-450-763-45646 (1-453)

QY	91	GAAGAGCCACTCTCTAGAAAGAGTTCTCGCGGTTTGTCTATCTTCCATCCAGTACCCT	150
DB	133	AspGluProLeuLeuGluAsnProArgPheValIlePheProIleGluTyrHis	152
QY	151	GATATTGGAAATGTATAACACAGGCACAGGCTTCTCTCGACAGCAGAGAGGTCGAC	210
DB	153	AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluValAsp	172
QY	211	TTATCAAGAGATCTCCCTCACTGGAAACAAGCTTAAACAGATGAGAGTACTCTCT	270
DB	173	LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer	192
QY	271	CACATCTTAGCCTTTTTCAGCCAGTGTGGAATTTGTAATCAAAATTTGGTGGAGCGC	330
DB	193	HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg	212
QY	331	TTTAGTCAGAGGTGCAGGTTCAGAGGCTCGCTGTTTCTATGCGTTTCAAAATTTCTCATC	390
DB	213	PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet	232
QY	391	GAGAATGTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAGATCCCAAG	450
DB	233	GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys	252
QY	451	AAAAGGAATTTTATTAATGCAATGAAACCATGCCCTATGTTAAGAAAAAAGCAGAT	510
DB	253	GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysLysAlaAsp	272
QY	511	TGGGCTTCGGATGAGATAGCAGATAAATCTACTTTTGGGAAAGAGTGGGCTTTT	570
DB	273	TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaAlaPhe	292
QY	571	GCTGCTGTAGAAGGAGTTTCTCTCAGGATCTTTTGTCTGTATATCTTGGCTAAAGAAG	630
DB	293	AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys	312
QY	631	AGAGTCTTATGCCAGGATCTACTTTTCCAATGAATCATCAGCAGAGATGAAGACTT	690
DB	313	ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu	332
QY	691	CAGTGTGACTTTCCTGCTGATGTTCATTAATTAAGCTTCAAGAGAGG	750
DB	333	HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg	352
QY	751	GTGAGGAGATCATTTGTCATGCTGCAAAATTTGACAGAGGTTTTTAACAGAGGCTTG	810
DB	353	ValArgGluIleIleIleAsnAlaValArgIleGluGlnGluPheThrGluAlaLeu	372
QY	811	CCAGTTGGCCTCATTTGGAATGAATTTGATGAAACAGTACATTTGAGTTTGTAGCT	870
DB	373	ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla	932

QY	871	GACAGATTACTTGTGGAACTTGGATTCTCAAAAGTTTTCAGGACAGAAAAATCCTTTGAT	930
DB	393	AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp	412
QY	931	TTTATGAAAAACATTTCTTTAGAGGAAAAACAATTTCTTTGAGAAACGAGTTTCAGAG	990
DB	413	PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu	432
QY	991	TATCAGCGTTTTCAGATTATGGCAACACACAGATAACGCTTTCACCTTGGATGCAGAT	1050
DB	433	TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp	452
QY	1051	TTT 1053	
DB	453	Phe 453	

RESULT 9
US-10-403-571-76
; Sequence 76, Application US/10403571
; Publication No. US20040068763A1
; GENERAL INFORMATION:
; APPLICANT: Hopkins, Nancy
; APPLICANT: Golling, Gregory
; APPLICANT: Amsterdam, Adam
; APPLICANT: Sun, Zhaoxia
; TITLE OF INVENTION: Developmental Mutations in Zebrafish
; FILE REFERENCE: 01997/539002
; CURRENT APPLICATION NUMBER: US/10/403,571
; CURRENT FILING DATE: 2003-03-25
; PRIOR APPLICATION NUMBER: US 60/368,760
; PRIOR FILING DATE: 2002-03-29
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 76
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Danio rerio
US-10-403-571-76

Alignment Scores:
Pred. No.: 1,09e-144 Length: 386
Score: 1425.50 Matches: 268
Percent Similarity: 89.41% Conservative: 36
Best Local Similarity: 78.82% Mismatches: 31
Query Match: 77.22% Indels: 5
DB: 15 Gaps: 1

US-10-698-228-2 (1-1053) x US-10-403-571-76 (1-386)

QY	49	AGATCATCTTCAGACACCAAGAAAGTAAATAAGTCA-----AATGAA	93
DB	47	ArgLysIlePheAspGluSerGluGlyGlnSerLysAlaLysLysGlyAlaValGluGlu	66
QY	94	GAGCCTCTCTAAGAAAGAGTCTCGCGGTTTCTCATCTTTCCAAATCCAGTACCTGAT	153
DB	67	GluProLeuLeuLysGluAsnProHisArgPheValIlePheProIleGlnTyrHisAsp	86
QY	154	ATTGGAAAAATGATAAACAGGACAGGCTTCTTCTGGACAGCAGAGAGTTCGACTTA	213
DB	87	IleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAspLeu	106
QY	214	TCAAAGATCTCCCTCACTGGAACACAGCTTAAACAGCATGAGAGTACTTCTCTCAT	273
DB	107	SerLysAspLeuGlnHisTrpAspSerLeuLysAspGluGluArgTyrPheIleSerHis	126
QY	274	ATCTTAGCCTTTTTCAGCGCAGTGTGAATTCGTAATGAAATTTTGGTGGAGCGCTTT	333
DB	127	ValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArgPhe	146
QY	334	AGTCAGAGGTGCAGGTTCCAGAGGCTCGCTGTTTCTATGGCTTTTCAAAATCTCATCG	393
DB	147	ThrGlnGluValGlnValThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMetGlu	166

```
QY 394 AATGTTCACTCAGAGATGACAGATTTGCTGCTAGACACTTACATCAGAGATCCCAAGAA 453
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
167 AsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspSerLysGlu 186
QY 454 AGGGAATTTTATTAATGCAATTGAAACCATGCCCTATGTAAGAAAAACAGATTGG 513
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
187 ArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysLysAlaAspTrp 206
QY 514 GCCTTGCGATGATGACAGATAGAAATCTACTTTTGGGGAAGAGTGGTGCCTTGGCT 573
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
207 AlaLeuAsnTrpIleCysAspLysAsnAlaArgTyrGlyGluArgValValAlaPheAla 226
QY 574 GCTGTAGAGAGGAGTTTCTCTCAGGATCTTTTGTGCTATATTTCTGGCTAAAGAGAGA 633
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
227 AlaValGluGlyLePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLysArg 246
QY 634 GGTCTTATGCCAGGACTCACTTTTCCAAATGATCATCAGCAGAGATGAAGACTTCAAC 693
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
247 GlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHis 266
QY 694 TGTGACTTGTCTGCTGATGTTTCCAAATCTAGTAAATAGCCCTTCAGAGAAAGGCTC 753
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
267 CysAspPheAlaCysLeuMetPheLysHisLeuIleAsnLysProSerGluGluThrVal 286
QY 754 AGGAGATCATTTGATGCTGCTCAAAATTTGACAGAGAGTTTAAACAGAAAGCTTGCCA 813
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
287 LysLysIleIleMetAsnAlaValGluIleGluGluPheLeuThrAspAlaLeuPro 306
QY 814 GTTGGCCCTCATTTGAATGAAATTTGATGAAACAGTACATGAGTTTGTAGCTGAC 873
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
307 ValLysLeuIleGlyMetAsnCysAspLeuMetLysGlnTyrIleGluPheValAlaAsp 326
QY 874 AGATTACTTGTGGAAGTCTCAAGGTTTTCAGGCGAAGAAATCTTTGATTTT 933
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
327 ArgLeuLeuLeuGluGlyPheAspLysValTyrArgValGluAsnProPheAspPhe 346
QY 934 ATGGAACCAATTTCTTAGAAGAAACAAATTTCTTTGAGAAACGAGTTTCAGAGTAT 993
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
347 MetGluAsnIleSerLeuGluGlyLysThrAsnPheGluLysArgValGlyGluTyr 366
QY 994 CAGCCTTTTGCAGTTATGCGAGAAACCAAGATAAGCTTCCACCTTGCATGAGATTTT 1053
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
367 GlnArgMetGlyValMetSerGlyThrThrAspAsnThrPheThrLeuAspAlaAspPhe 386

RESULT 10
US-11-097-143-19182
; Sequence 19182, Application US/11097143
; Publication No. US2005020858A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
```

```
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 19182
; LENGTH: 393
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-19182

Alignment Scores:
Pred. No.: 2,67e-136 Length: 393
Score: 1348.00 Matches: 252
Percent Similarity: 85.92% Conservative: 41
Best Local Similarity: 73.90% Mismatches: 48
Query Match: 73.02% Indels: 0
DB: 20 Gaps: 0

US-10-698-228-2 (1-1053) x US-11-097-143-19182 (1-393)
QY 31 GGCTCGATCAGGATGAGAGATCTTTCAGACACCAACGAAAGTGAATAAAGTCAAT 90
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
53 GlyIleGlyLysSerAlaAsnSerLeuMetGluLysSerValThrProPheAspProSer 72
QY 91 GAAGAGCCACTCTTAAGAAAGAGTTCTCGCCGGTTTGTCTTCCATCCAGTACCCT 150
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
73 LeuGluProLeuLeuArgGluAsnProArgArgPheValIlePheProIleGlnTyrHis 92
QY 151 GATATTGGAAATGTATAACAGGCACAGCTTCTCTGACAGCAGAGAGGTCCGAC 210
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
93 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrValGluGluValAsp 112
QY 211 TTATCAAAAGATCTCCCTCACTGCAACCAAGCTTAAAGCAGATGAGAACTTCTCATCTCT 270
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
113 LeuSerLysAspLeuThrAspTrpHisArgLeuLysAspAspGluArgHisPheIleSer 132
QY 271 CACATCTTAGCCTTTTTCAGCCAGCTGATGGAATTTGAAATGAAATTTTGGTGAGCGC 330
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
133 HisValLeuAlaPhePheAlaIleAspGlyIleValAsnGluAsnLeuValGluArg 152
QY 331 TTATGTCAGAGCTGCGAGTTCCAGAGGCTCGCTGTTCTATGCTTCAATTTCTCATC 390
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
153 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 172
QY 391 GAGAATGTTCACTCAGAGATGTACAGTTGCTGTAGTACACTTACATCAGAGATCCCCAAG 450
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
173 GluAsnValHisSerGluMetTyrSerValLeuIleAspThrTyrIleArgAspProHis 192
QY 451 AAAAGGAATTTTATTTAATGCAATTTGAAACCATGCCCTATGTTAAGAAAAAGCAGAT 510
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
193 GlnArgGluTyrLeuPheAsnAlaIleGluThrMetProAlaValLysArgLysAlaAsp 212
QY 511 TGGGCTTGGATGATGATAGAGAAATCTACTTTTGGGGAAGAGTGGTGGCCTTT 570
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
213 TrpAlaLeuSerTrpIleSerSerLysSerAlaAsnPheGlyGluArgIleAlaPhe 232
QY 571 GCTGCTGTAGAAGAGTTTCTCTCAGGATCTTTTCTGCTGTATATTCTGCTGCTAAAGAG 630
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
233 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 252
QY 631 AGAGTCTTATGCGCAGGACTCCTCTTTTCCAACTCACTCAGCAGAGATGAAGGACTT 690
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
253 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 272
QY 691 CACTGTGACTTTCCTGCTGATGTTCCAAATCTTCAATCTAGTAATAAGCCTTCAGAGAAAGG 750
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
273 HisCysAspPheAlaValLeuMetPheGlnHisLeuValGlnArgProLysArgGluArg 292
QY 751 GTCAGGAGATCATTTGTTGATGCTGCAAAATTCAGCAGGAGTTTAAACAGAGCCTTG 810
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
293 IleIleGluIleIleArgAspAlaValAlaIleGluGlnGluPheLeuThrAspAlaLeu 312
QY 811 CCAGTTGGCCTCATTTGGAATGAAATTTGATGAAACAGTACATTGAGTTGTAGCT 870
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
313 ProValAsnLeuIleGlyMetAsnCysAspLeuMetSerGlnTyrIleGluPheValAla 332
```

QY 871 GACAGATTACTTGTGGAACCTTGATCTCAAGAGTTTTCAGGCAGAAATCCTTTTGAT 930
 DB 333 AspArgLeuLeuValGluLeuGlyValGlyLysIleTyrAsnThrLysAsnProPheAsn 352
 QY 931 TTTATGGAAAACATTTCTTAGAGGAAAAAACAATTTCTTTCAGAAAACGAGTTTCAGAG 990
 DB 353 PheMetGluMetIleSerLeuAspGlyLysThrAsnPhePheGluLysLysValGlyGlu 372
 QY 991 TATCAGCGTTTTCAGTTATGCGAGAAACACACAGATAACGCTCTTCACCTTGATGATGAGAT 1050
 DB 373 TyrGlnArgMetGlyValValSerAsnProLeuAspAsnValPheThrLeuAspAlaAsp 392
 QY 1051 TTT 1053
 DB 393 Phe 393
 RESULT 11
 US-10-369-493-4122
 ; Sequence 4122, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; PRIOR FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 4122
 ; LENGTH: 430
 ; TYPE: PRT
 ; ORGANISM: Neurospora crassa
 US-10-369-493-4122
 Alignment Scores:
 Pred. No.: 8,12e-125 Length: 430
 Score: 1242.00 Matches: 243
 Percent Similarity: 76.03% Conservative: 52
 Best Local Similarity: 62.63% Mismatches: 47
 Query Match: 67.28% Indels: 46
 DB: 15 Gaps: 6
 US-10-698-228-2 (1-1053) x US-10-369-493-4122 (1-430)
 QY 28 GCCGGGCTGGATCAGATGAGATCATCTTCAGACACCAACGAAAGTGAA----- 78
 DB 43 AlaGluIleAspAlaGluHisAsnAlaAsnLysLysAlaAlaGluAlaLysLysMetAla 62
 QY 79 -----ATAAGTCAATGAA-----GAGCCACTCTTAAGAAAGAGTTCTCGCGGTTT 126
 DB 63 ProThrLeuLysProGluGluAlaAsnGluProLeuLeuThrGluAsnProGlnArgPhe 82
 QY 127 GTCATCTTTTCAATCCAGTACCTCAT----- 153
 DB 83 ValLeuPheProIleLysTyrHisGluValCysHisAlaSerGlnThrPheGlnThrThr 102
 QY 154 -----ATTGGAAAATGATATAA 171
 DB 103 ArgLysValValHisGlyLeuIleGlyTyrProHisTyrArgIleTyrGlnMetTyrLys 122
 QY 172 CAGGCACAGGTTCTTCTGGACGACGAGAGGTGCGACTTATCAAGATCTCCCTCAC 231
 DB 123 LysAlaGluAlaSerPheThrAlaGluGluIleAspLeuSerLysAspLeuHisAsp 142
 QY 232 TGG---AACAAAGCTTAAGCAGATGAGAAGTACTCTCTCATCTCTCATCTTACGCTTTT 288

Db 143 TrpAsnAsnArgLeuAsnAspAspGluLysPhePheIleSerHisIleLeuAlaPhePhe 162
 QY 289 GCAGCCAGTCATGAAATTTGTAATAATTTGTGGAGCGCTTTAGTCAGAGGTCGAG 348
 Db 163 AlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerGlyGluValGln 182
 QY 349 GTTCAGAGGCTCGCTGTTCTATCGGCTTTCAAAATTTCTCATCGAATAATTTCACTCAGAG 408
 Db 183 IleProGluAlaArgCysPheTyrGlyPheGlnIleMetMetGluAsnIleHisSerGlu 202
 QY 409 ATGTACAGTTGCTGTAGACACCTATCATCAGAGATCCCAAGAAAGGAAATTTTATTT 468
 Db 203 ThrTyrSerLeuLeuIleAspThrTyrIleLysGluProSerGlnArgThrTyrLeuPhe 222
 QY 469 AATGCAATTCGAACCATGCGCTATGTTAAGAAAAGACAGATTTGGCTTTCGATGGATA 528
 Db 223 AsnAlaIleAspThrIleProCysIleArgLysLysAlaAspTrpAlaLeuArgTrpIle 242
 QY 529 GCAGATAGAAAATCTACTTTTGGGGAAGAGTGTGGCTTTTGTCTGTGTAGAAGGAGTT 588
 Db 243 ThrAspLysSerSerThrPheAlaGlnArgLeuValAlaPheAlaAlaValGluGlyIle 262
 QY 589 TTCTTCTCAGGATCTTTTGTCTATATTTCTGGCTAAAGAGAGAGGCTTTATGCCAGGA 648
 Db 263 PhePheSerGlyAlaPheAlaSerIlePheTrpLeuLysLysArgGlyLeuMetProGly 282
 QY 649 CTCACTTTTCCATGAACTCATCAGCAGATGAGGACTTCTACTGTGACTTGTCTGCTTGC 708
 Db 283 LeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisThrAspPheAlaCys 302
 QY 709 CTGATGTTTCCAAATCTAGTAAATAAGCCTTTCAGAAAGAGGTCAGGAGATCATTTGT 768
 Db 303 LeuLeuPheSerHisLeuAsnAsnArgProSerLysGlnLeuIleGlnIleVal 322
 QY 769 GATGCTCTCAAAATTCAGCAGGAGCTTTTAAACAGAGCCTTGCAGGTTGCGCTCATTTGA 828
 Db 323 AspAlaValArgIleGluGlnGluPheLeuThrGluAlaLeuProCysAlaLeuGly 342
 QY 829 ATGAATTGCAATTTGTAGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGAA 888
 Db 343 MetAsnAlaAspLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValAla 362
 QY 889 CTTGGATCTCAAGGTTTTCAGGCAGAAATCCTTTGATTTTATCGAAAACATTTCT 948
 Db 363 LeuGlyAsnGluLysIleTyrArgSerThrAsnProPheAspPheMetGluAsnIleSer 382
 QY 949 TTAGAAAGGAAAAACAAATTTCTTTCAGAAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTT 1008
 Db 383 LeuGlyGlyLysThrAsnPhePheGluLysArgValGlyAspTyrGlnLysAlaGlyVal 402
 QY 1009 ATG-----GCAGAAACACAGATAACGTC----- 1032
 Db 403 MetAsnSerThrLysLysAlaAspAlaAspAlaGluValAlaLysAsnGluAsnGlyGly 422
 QY 1033 ---TTCACCTTGGATCGAGATTTT 1053
 Db 423 AspPheThrPheAspGluAspPhe 430
 RESULT 12
 US-10-128-714-8245
 ; Sequence 8245, Application US/10128714
 ; Publication No. US20030119013A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Jiang, Bo
 ; APPLICANT: Hu, Weng
 ; APPLICANT: Tishkoff, Daniel
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Broshkin, Alexey M
 ; APPLICANT: Lemieux, Sebastien M
 ; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
 ; FILE REFERENCE: 10182-018-999
 ; CURRENT APPLICATION NUMBER: US/10/128,714

; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8245
; LENGTH: 405
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-128-714-8245

Alignment Scores:
Pred. No.: 158e-123 Length: 405
Score: 1230.00 Matches: 234
Percent Similarity: 81.95% Conservative: 52
Best Local Similarity: 67.05% Mismatches: 49
Query Match: 66.63% Indels: 14
DB: 14 Gaps: 3

US-10-698-228-2 (1-1053) x US-10-128-714-8245 (1-405)

QY 46 GAGAGATCATCTTCAGACACCAAGAAAGTGAAATCAATCAAGAGCCACTCTTA 105
DB 58 GluAlaSerLysValAlaGlnSerIleLysGluLeuGluAlaAsn---GluProLeuLeu 76
QY 106 AGAAGAGTTCTCGCGGTTTCTCATCTTTCATCCAGTACCTCGATATTGGAAAATG 165
DB 77 GlnGluAsnProHisArgPheValLeuPheProIleLysTyHisGluIleTrpGlnMet 96
QY 166 TATAACAGGACAGGCTTCCTCTGGACAGACAGAGGTGACTTATCAAGATCTC 225
DB 97 TyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluIleAspLysSerLysAspLeu 116
QY 226 CTCACTGG---AACAGCTTAAAGCAGATGAGAAGTACTTCTCATCTCAGATCTTAGCC 282
DB 117 HisAspTrpAsnAsnArgLeuAsnAspAspGluAArgTyrPheIleSerHisValLeuAla 136
QY 283 TTTTTCAGCCAGTGATGGAATTGTAATGAAAATTTCGTGAGCGCTTTAGTCAGAG 342
DB 137 PhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuLeuGluArgPheSerAsnGlu 156
QY 343 GTGAGGTTCCAGAGCTCGCTGTTTCTATGCTTCAATTTCTCATCGAGATGTTTAC 402
DB 157 ValGlnValProGluAlaArgCysPheTyrGlyPheGlnIleMetIleGluAsnIleHis 176
QY 403 TCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAAATT 462
DB 177 SerGluThrTyrSerLeuLeuIleAspThrTyrIleLysGluProLysGlnArgThrTyr 196
QY 463 TTATTATGCAATTGAAACCACTGCCCTATGTTAAAGAAAAGCAGATTTGGCCCTTGGCA 522
DB 197 LeuPheAspAlaIleAspThrValProCysIleArgLysLysAlaGlnTrpAlaMetArg 216
QY 523 TGGATAGCAGATAGAAAATCTACTTTTGGGAAAAGAGTGGTGGCTTCTGCTGTAGAA 582
DB 217 TrpIleSerAspLysGluSerThrPheAlaGlnArgLeuValAlaPheAlaAlaValGlu 236
QY 583 GGAGTTTCTCTCAGGATCTTTTCTGCTATATCTTGCTAATGCTAAGAGAGAGTCTTAG 642
DB 237 GlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysArgGlyLeuMet 256
QY 643 CCAGGACTCAGCTTTTTCGAATCAATCATCAGCAGAGATGAAGGACTTCACTGTGACTTT 702
DB 257 ProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisThrAspPhe 276

QY 703 GCTTGCTGATGTTCCAAATCTTAGTAAATAAAGCCTTTCAAGAAAGAGGTCCAGGAGATC 762
DB 277 AlaCysLeuLeuPheSerHisLeuAsnHisArgProAspLysLysValIleGluAspIle 296
QY 763 ATTGTTGATGCTGTCAAAATTGAGCAGGAGTCTTTTAAACAGAGCCTTGCAGTTGGCTC 822
DB 297 IleValGluAlaValAlaIleGluGlnGluPheLeuThrAspAlaLeuProValAlaLeu 316
QY 823 ATTGGAATGAATTGTCATTTTGATGAAACAGCTACATTCAGTTGTAGCTGACAGATTACTT 882
DB 317 LeuGlyMetAsnSerLysLeuMetCysGlnTyrIleGluPheValAlaAspArgLeuLeu 336
QY 883 GTGGAACCTTGGATTCTCAAGGTTTTCAGCGCAAAATCCTTTTGATTTTATGCAAAAC 942
DB 337 ValAlaLeuGlyAsnLysLysTyrPheAsnSerProAsnProPheAspPheMetGluSer 356
QY 943 ATTCTTTTGAAGGAAAAACAAATTTCTTTTGAGAAACGAGTTTCAGAGTATCAGCGTTT 1002
DB 357 IleSerLeuAlaGlyLysThrAsnPheGluLysArgValGlyAspTyrGlnLysAla 376
QY 1003 GCAGTTTATGCGCAAAACC-----ACAGAT 1026
DB 377 GlyValMetAlaSerThrLysLysAspAlaSerGlnAspThrAlaLysAspAlaAsnAsp 396
QY 1027 AACGTTCTCACCTTGGATGTCAGATTTT 1053
DB 397 GlyGlyLeuCysPheAspGluAspPhe 405

RESULT 13
US-10-369-493-5686
; Sequence 5686, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 5686
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-5686

Alignment Scores:
Pred. No.: 8.3e-122 Length: 381
Score: 1214.00 Matches: 231
Percent Similarity: 81.10% Conservative: 48
Best Local Similarity: 67.15% Mismatches: 61
Query Match: 65.76% Indels: 4
DB: 15 Gaps: 2

US-10-698-228-2 (1-1053) x US-10-369-493-5686 (1-381)

QY 34 CTGGATCAGATGAGATCATCTTCAGAC-----ACCAACGAAAGTGAATAAG--- 84
DB 38 ValAspGlnThrLysAlaAlaSerAlaGluGluThrAsnAsnGluSerGluValAsnGlu 57
QY 85 ---TCAATCAAGAGCCACTCTTAAGAAAGAGTTCCTCGCGGTTTGTTCATCTTCCCAATC 141
DB 58 LeuAspAlaAspGluProMetLeuGlnAspLeuAspAsnArgPheValIlePheProLeu 77
QY 142 CAGTACCTCATATTGGAATAATGTATAAAACAGGACAGGCTTCTCTTCGACAGACAGAA 201
DB 78 LysHisHisAspIleTrpAsnPheTyrLysLysAlaValAlaSerPheTrpThrValGlu 97

```
QY 202 GAGGTGCACTTATCAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAAGTAC 261
|||
Db 98 GluValAspLeuGlyLysAspMetAsnAspTrpGluLysMetAsnGlyAspGluGlnTyr 117
|||
QY 262 TTCATCTCTCATCTTAGCTTTTTCAGCCAGTGTGAATTTGTAATGAATAATTTG 321
|||
Db 118 PheIleSerArgIleLeuAlaPhePheAlaIleAspGlyIleValAsnGluAsnLeu 137
|||
QY 322 GTGGAGCGCTTAGTCAGGAGTGCAGGTTCCAGAGGCTCGTGTCTTCTATGCTTCAA 381
|||
Db 138 CysGluArgPheSerAsnGluValGlnValSerGluAlaArgPheTyrGlyPheGln 157
|||
QY 382 ATTCTCATCGAATGTTTCACTCAGAGATGTACAGTTTCTGATGATACACTTACATCAGA 441
|||
Db 158 IleAlaIleGluAsnIleHisSerGluMetTyrSerLysLeuIleGluThrTyrIleArg 177
|||
QY 442 GATCCCAAGAAAGGAATTTTATTATTAATCAATGAAACCATGCCCTATGTTAAGAAA 501
|||
Db 178 AspGluThrGluArgAsnThrLeuPheAsnAlaValAspGluPheGluPheIleLys 197
|||
QY 502 AAAGCAGATTGGCTTGGATGGATAGCAGATAGATAAATCTACTTTTGGGGAAGAGTG 561
|||
Db 198 LysAlaAspTrpAlaLeuArgTrpIleSerAspLysLysAlaSerPheAlaGluArgLeu 217
|||
QY 562 GTGGCTTGTCTGTAGAGGATTTTCTTCTCAGGATCTTTTCTGCTATATTTCTGG 621
|||
Db 218 IleAlaPheAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrp 237
|||
QY 622 CTAAAGAAGAGAGGTCTTATGTCAGGACTCACTTTTCCAAATCACTATCAGCAGAGAT 681
|||
Db 238 LeuLysLysArgGlyLeuMetProGlyLeuThrHisSerAsnGluLeuIleSerArgAsp 257
|||
QY 682 GAAGACTTCACTGTGACTTCTGCTGCTGATGTTTCCAAATCACTAGTAATTAAGCCTTCA 741
|||
Db 258 GluGlyLeuHisArgAspPheAlaCysLeuLeuTyrSerLysLeuGlnLysLysLeuThr 277
|||
QY 742 GAAGAAGGTCAGGAGATCACTTGTGATGCTGCTCAAAATTTGACAGAGTTTAAACA 801
|||
Db 278 GlnGlnArgIleAspIleIleLysAspAlaValAlaIleGluGlnGluPheLeuThr 297
|||
QY 802 GAAGCTTGCCTGCTCACTTGAATGAATTTGATTTTCAATGAAACAGTACATTTGAG 861
|||
Db 298 GluAlaLeuProValAspMetIleGlyMetAsnCysArgLeuMetSerGlnTyrIleGlu 317
|||
QY 862 TTTGAGTGCAGATTTACTTGTGAACCTTGATTTCTCAAGGTTTTCAGGCAGAAAT 921
|||
Db 318 PheValAlaAspHisLeuLeuValGluLeuGlyCysAspLysLeuTyrLysSerLysAsn 337
|||
QY 922 CTTTTGATTTTATGAAAAAATTTCTTTAGAGGAAAAAATAATTTCTTTGAGAAACGA 981
|||
Db 338 ProPheAspPheMetGluAsnIleSerIleAspGlyLysThrAsnPhePheGluLysArg 357
|||
QY 982 GTTTCAGAGTATCAGCGTTTTCAGTTATGTCAGAAACCCACAGATAAATCTTCACTTGG 1041
|||
Db 358 ValSerGluTyrGlnArgProGlyValMetValAsnGluAlaGluArgGlnPheAspLeu 377
|||
QY 1042 GATCAGATTTT 1053
|||
Db 378 GluAlaAspPhe 381
|||
```

RESULT 14

```
US-10-128-714-3245
; Sequence 3245, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wengqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
```

```
; TITLE OF INVENTION: Methods of Use
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3245
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-128-714-3245
```

Alignment Scores:

```
Pred. No.: 1,76e-119 Length: 381
Score: 1192.50 Matches: 229
Percent Similarity: 80.23% Conservative: 51
Best Local Similarity: 65.62% Mismatches: 50
Query Match: 64.60% Indels: 19
DB: 14 Gaps: 4
```

US-10-698-228-2 (1-1053) x US-10-128-714-3245 (1-381)

```
QY 46 GAGAGATCATCTTCAGACACCAAGAAAGTGAATAAGTCAATGAAGACCCACTCTCTA 105
|||
Db 39 GluAlaSerLysValAlaGlnSerIleLysGluLeuAlaAsn--GluProLeuLeu 57
|||
QY 106 AGAAAGAGTTCTCGCGGTTTGTCACTTTCCATCCAGTACCCTGTATATTGGAAATG 165
|||
Db 58 GlnGluAsnProHisArgPheValLeuPheProIleLysTyrHisGluIleTrpGlnMet 77
|||
QY 166 TATAAAGCAGGCACAGGCTTCTTCTGACACGACGAAAGAGTTCGACTTATCAAGGATCTC 225
|||
Db 78 TyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluIleAspLeuSerLysAspLeu 97
|||
QY 226 CCTCACTGG--AACAGCTTAAAGCAGATGAGAAGTACTTCACTCTCTCACTTACGCC 282
|||
Db 98 HisAspTrpAsnAsnArgLeuAsnAspAspGluArgTyrPheIleSerHisValLeuAla 117
|||
QY 283 TTTTTCGACCCAGTGTGAATTTGAATGAAATTTGGTGGAGCGCTTTAGTCAGGAG 342
|||
Db 118 PhePheAlaAlaSerAspGlyIleValAsnGluAsnPhe-----AsnGlu 132
|||
QY 343 GTGCAGGTTCACAGGCTCGCTGTTTCTATGGCTTTTCAAAATTTCTCATCGAATGTTTCCAC 402
|||
Db 133 ValGlnValProGluAlaArgCysPheTyrGlyPheGlnIleMetIleGluAsnIleHis 152
|||
QY 403 TCAGAGATGTACAGTTTGTGTAGACACTTATCATCAGAGATCCCAAGAAAGGAATTT 462
|||
Db 153 SerGluThrTyrSerLeuLeuIleAspThrTyrIleLysGluProLysGlnArgThrTyr 172
|||
QY 463 TTATTTAATGCAATTTGAAACCATGCCCTATGTTAAGNAAAAGCAGATTTGGCCTTGGCA 522
|||
Db 173 LeuPheAspAlaIleAspThrIleProCysIleArgLysLysAlaGlnTrpAlaMetArg 192
|||
QY 523 TGGATAGCAGATAGAAAAATCTACTTTTGGGGAAGAGTGGTGGCTTTTGGCTGTAGAA 582
|||
Db 193 TrpIleSerAspLysGluSerThrPheAlaGlnArgLeuValAlaPheAlaValGlu 212
|||
QY 583 GGAGTTTCTCTCAGGATCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGCTTTATG 642
|||
Db 213 GlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLysArgGlyLeuMet 232
|||
QY 643 CCAGGACTCACTTTTCCAAATGAATCATCATCAGCAGATGAAGGACTTCCTGACTGACTTT 702
|||
```

Db 233 ProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisThrAspPhe 252
QY GCTTCCCTGATCTTCCAACTACTAGTAATAAGCCTTCAGAGAGAAAGGTCAGGAGATC 762
Db 253 AlaCysLeuLeuPheSerHisLeuAsnHisArgProAspLysLysValIleGluAspIle 272
QY ATTGTGTATGCTGTCAAAATTCAGCAGGAGTATTTAAACAGAGAGCCTTCAGTTCGCCCTC 822
Db 273 IleValGluAlaValAlaIleGluGlnPheLeuThrAspAlaLeuProValAlaLeu 292
QY ATTGAATGAATTCATTTTGATGAACAGTACATGATGTTGTAGTCTCAGATTAATTT 882
Db 293 LeuGlyMetAsnSerLysLeuMetCysGlnIleGluPheValAlaAspArgLeuLeu 312
QY GTGGAACTTGGATCTCAAGGTTTTCAGGCAGAGAAATCTTTGATTTATGGAAAC 942
Db 313 ValAlaLeuGlyAsnLysLysTyrPheAsnSerProAsnProPheAspPheMetGluSer 332
QY ATTTCTTTAGAGAGAAACAAATTTCTTTGAGAACGAGTTTCAGAGTATCAGCGTTT 1002
Db 333 IleSerLeuAlaGlyLysThrAsnPhePheGluLysArgValGlyAspTyrGlnLysAla 352
QY GCAGTTATGGCAGAAAC-----ACAGAT 1026
Db 353 GlyValMetAlaSerThrLysLysAspAlaSerGlnAspThrAlaLysAspAlaAsnAsp 372
QY 1027 AACGCTTCACCTTCGATCCAGATTTT 1053
Db 373 GlyGlyLeuCysPheAspGluAspPhe 381

RESULT 15

US-10-032-585-7537
; Sequence 7537, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7537
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-032-585-7537

Alignment Scores:
Pred. No.: 2,05e-119 Length: 413
Score: 1192.00 Matches: 225
Percent Similarity: 80.17% Conservative: 58
Best Local Similarity: 63.74% Mismatches: 64
Query Match: 64.57% Indels: 6
DB: 14 Gaps: 3

US-10-698-228-2 (1-1053) x US-10-032-585-7537 (1-413)

QY 13 GAAAGCGGAGCGCGCGGTGGATCAGATGAGATCATCTTCAGACACCAACGAA 72
Db 61 GluLysAspSerLeuLysLysHisGlnAspPheLeuAlaLysHisLysValHisArg 80
QY 73 AGTGAATAAAG-----TCAATGAGAGCCACTCTTAAGAAAGAGTTCTCGCCGTTT 126
Db 81 HisLysLeuLysGlnLeuGluAlaGluProLeuValGluAsnLysArgArgTyr 100
QY 127 GTCATCTTTCCCAATCAGTACCTGATATTTGAAATCGTATAAACAGGCACAGGCTTC 186
Db 101 ValMetPheProIleArgTyrHisGluIleTyrAsnPheTyrLysLysAlaGluAlaSer 120

QY 187 TTCTGCACGACAGAGGTCGACTTATCAAAGGATCTCCCTCACTGG---AACAAAGCTT 243
Db 121 PheTyrAlaGluGluIleAspLeuSerLysAspLeuAspTyrPheAsnLysLeu 140
QY 244 AAGCAGATGAGAAGTACTCTCTCTCATCTTACGCTTTTGGCAGCAGCATGATGA 303
Db 141 AsnGluAsnGluArgTyrPheIleSerArgValLeuAlaPhePheAlaAlaSerAspGly 160
QY 304 ATTGTAATAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCTCGC 363
Db 161 IleValGlyGluAsnLeuIleGluAsnPheSerThrGluValGlnLeuProGluAlaLys 180
QY 364 TGTTTCTATGCTTTCAAATTCATCGAGAATGTTCTCACTCAGAGATGTCAGTTTGTCTG 423
Db 181 SerPheTyrGlyPheGlnIleMetMetGluAsnIleHisSerGluThrTyrSerLeuLeu 200
QY 424 ATAGACATTTACATCAGAGATCCCAAGAAAGGAATTTTATTAATGCAATTTGAAACC 483
Db 201 IleGluThrTyrIleLysAspProGlnGluAlaAspTyrLeuPheAsnAlaIleAlaAsn 220
QY 484 ATGCCCTATGTTAAGAAAGAGAGATTTGGGCTTCGATGGATAGCAGATAGAAAATCT 543
Db 221 IleProCysIleGlnLysLysAlaAspTyrPheIleLysTrpIleGlnAspAspGluAla 240
QY 544 ACTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAAGAGTTTCTTCTCAGGATCT 603
Db 241 LeuTyrGlyGluArgLeuValAlaPheAlaAlaValGluGlyIlePhePheSerGlySer 260
QY 604 TTTGCTGCTATATCTGGCTTAAAGAGAGAGGCTTTATGCCAGGACTCACATTTTCCAAT 663
Db 261 PheAlaSerIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSerAsn 280
QY 664 GAATCATCAGCAGATGAGAGCTTCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 723
Db 281 GluLeuIleCysArgAspGluGlyLeuHisThrAspPheAlaCysLeuLeuPheAlaHis 300
QY 724 TTAGTAAATAAGCCTTCAGAAAGAGGCTCAGGAGAGATCATTTGTGATGCTGTCAAAAT 783
Db 301 LeuGlnAsnArgProSerProGluIleValGluArgIleIleThrGluAlaValAspVal 320
QY 784 GAGCAGAGTATTTTAAACAGAGGCTTCCAGTTCGCTTCATTTGGAATGAATTTGTTG 843
Db 321 GluLysGluTyrPheThrAspAlaLeuProValSerLeuLeuGlyMetAsnCysAsnLeu 340
QY 844 ATGAAACAGTACATTTGAGTTTGTAGCTGACAGATTAATTTGTGGAACCTTCTCAAG 903
Db 341 MetCysGlnTyrIleGluPheValAlaAspArgLeuLeuAlaLeuGlyAsnLysLys 360
QY 904 GTTTTTCAGGCAGCAAAATCTTTTATGTTTATGAAACATTTCTTTAGAAAGGAAAAACA 963
Db 361 ValTyrAsnValThrAsnProPheAspPheMetGluAsnIleSerLeuAlaGlyLysThr 380
QY 964 AATTTCTTTCAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGGTTATGGCAGAAACCA 1023
Db 381 AsnPhePheGluLysArgValSerAspTyrGlnLysAlaGlyValMetAlaLysThrGlu 400
QY 1024 GAT-----AACGCTTCTCAGCTTCGATGTCAGATTTT 1053
Db 401 AsnLysGluAlaAspAlaPheThrPheAspGluAspPhe 413

Search completed: October 30, 2005, 06:55:05
Job time : 77.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2005, 21:42:15 ; Search time 3258 Seconds
(without alignments)
15660.940 Million cell upd.

Title: US-10-698-228-12

Perfect score:

Sequence: 1 atggcgaccggaaggcc.....tcacctggatgcagatttt 1053

Scoring table: IDENTITY NUC

Scoring cubic: IDENTITY_NOC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%

Fast-processing: Minimum Match 0%
Maximum Match 100%

Maximum Match 100%
Listing first 45 summaries

Database :

Database : GenEmbl:★

```
package :
  generate :
    1: qb ba:*
```

2: qb_htq:*

3: qb_in:*

4: gb_om: *

5: gb ov: *

6: gb_pat:*

7: gb_ph:*

8: gbpl:*

9: 9b pr:*

```
10: gb_ro:*
```

```
11: gb_sts:*
```

```
12: gb_sy:*
```

```
13: gb_un:*
```

```
14: gb_vi:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match	Length	DB	ID	Description
1	1053	100.0	1053	6	AR454877	AR454877	Sequence
2	1053	100.0	1053	6	BD064773	BD064773	Novel pro
3	1053	100.0	1053	6	BD093086	BD093086	Novel pro
4	1053	100.0	1081	6	AR454870	AR454870	Sequence
5	1053	100.0	1081	6	BD064766	BD064766	Novel pro
6	1053	100.0	1081	6	BD093079	BD093079	Novel pro
7	1051.4	99.8	1051	6	AR454868	AR454868	Sequence
8	1051.4	99.8	1053	6	BD064764	BD064764	Novel pro
9	1051.4	99.8	1053	6	BD093077	BD093077	Novel pro
10	1051.4	99.8	1056	6	CQ714252	CQ714252	Sequence
11	1051.4	99.8	1056	9	AB166671	AB166671	Homo sapi
12	1051.4	99.8	4955	6	AR454869	AR454869	Sequence
13	1051.4	99.8	4955	6	BD064765	BD064765	Novel pro
14	1051.4	99.8	4955	6	BD093078	BD093078	Novel pro
15	1051.4	99.8	4955	9	AB038063	AB038063	Homo sapi
16	1048.2	99.5	1601	6	BD156916	BD156916	Primer for
17	1048.2	99.5	1601	6	AX877905	AX877905	Sequence
18	946.8	95.5	1601	9	AK001965	AK001965	Homo sapi
19	898.2	85.2	4532	10	BC058103	BC058103	Mus mus

ALIGNMENTS

RESULT 1	AR454877	AR454877	1053 bp	DNA	linear	PAT 20-FEB-2004
LOCUS	DEFINITION	Sequence 12 from patent US 6682917.				
ACCESSION	AR454877					
VERSION	AR454877.1	GI:42688832				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1 (bases 1 to 1053)					
AUTHORS	Nakamura,Y., Arakawa,H. and Tanaka,H.					
TITLE	Protein having a ribonucleotide Reductase activity and a DNA thereof					
JOURNAL	Patent: US 6682917-A	12 27-JAN-2004;				
FEATURES	Location/Qualifiers					
source	1..1053					
	/organism="unknown"					
	/mol_type="genomic DNA"					
ORIGIN						
Query Match	100.0%;	Score 1053;	DB 6;	Length 1053;		
Best Local Similarity	100.0%;	Pred. No. 6.4e-255;				
Matches 1053;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;		
Qy	1	ATGGGGCACCCGGAAAGCGCGGCGGGCTGGATCAGGATGAGAGATCATCTTCA	60			
Db	1	ATGGGGCACCCGGAAAGCGCGGCGGGCTGGATCAGGATGAGAGATCATCTTCA	60			
Qy	61	GACACCAAGAAAGTGAATAAAGTCAATGAAGGCCACTCTTAAGAAAGATTCTCGC	120			
Db	61	GACACCAAGAAAGTGAATAAAGTCAATGAAGGCCACTCTTAAGAAAGATTCTCGC	120			
Qy	121	CGGTTTGTCATCTTTCCAAATCCAGTACCTCGATATTGGAAAAATGTATAAACAGGCACAG	180			
Db	121	CGGTTTGTCATCTTTCCAAATCCAGTACCTCGATATTGGAAAAATGTATAAACAGGCACAG	180			
Qy	181	GCTTCCTCTTGACAGCAGAGAGGTTGACTTATCAAGGATCTCCCTCAGTGGAAACAAG	240			
Db	181	GCTTCCTCTTGACAGCAGAGAGGTTGACTTATCAAGGATCTCCCTCAGTGGAAACAAG	240			
Qy	241	CTTAAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTAGCCCTTTTTCGAGCCAGTGAT	300			


```
VERSION AR454870.1 GI:42688825
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 1081)
AUTHORS Nakamura,Y., Arakawa,H. and Tanaka,H.
TITLE Protein having a ribonucleotide Reductase activity and a DNA
thereof
JOURNAL Patent: US 6682917-A 4 27-JAN-2004;
FEATURES
source Location/Qualifiers
1..1081
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 1053; DB 6; Length 1081;
Best Local Similarity 100.0%; Pred. No. 6.4e-255;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGCGACCGGAAAGCGGCGGCGGCTGGATCAGATCAGAGATCATCTTCA 60
DB 20 ATGGGCGACCGGAAAGCGGCGGCGGCTGGATCAGATCAGAGATCATCTTCA 79
QY 61 GACACCAACGAAAGTGAAATAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
DB 80 GACACCAACGAAAGTGAAATAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 139
QY 121 CGTTTGTGATCTTCCCAATCGATACCTGATATTTGGAATGATATTAACAGGACAG 180
DB 140 CGTTTGTGATCTTCCCAATCGATACCTGATATTTGGAATGATATTAACAGGACAG 199
QY 181 GCTTCTCTCTGACAGCAGAGAGTTGACTATCAAGAGATCTCTCACTGGAACAAG 240
DB 200 GCTTCTCTCTGACAGCAGAGAGTTGACTATCAAGAGATCTCTCACTGGAACAAG 259
QY 241 CTTAAGCAGATGAGAAGTACTTCTCTCACTCTTACATCTTACAGCTTTTTCGAGCAGTGAT 300
DB 260 CTTAAGCAGATGAGAAGTACTTCTCTCACTCTTACATCTTACAGCTTTTTCGAGCAGTGAT 319
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
DB 320 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 379
QY 361 CGCTGTTTCTATGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGATACAGTTTG 420
DB 380 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGATACAGTTTG 439
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 480
DB 440 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 499
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGATGAGATAGAGATAGAAAA 540
DB 500 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGATGAGATAGAGATAGAAAA 559
QY 541 TCTACTTTTGGGAAAGAGTGGGCTTTGCTGCTGATAGAGAGTTCCTCTCAGGA 600
DB 560 TCTACTTTTGGGAAAGAGTGGGCTTTGCTGCTGATAGAGAGTTCCTCTCAGGA 619
QY 601 TCTTTTGTGCTATATTCTGGCTTAAAGAAAGAGAGGTCTTATGCGAGACTCACATTTTTC 660
DB 620 TCTTTTGTGCTATATTCTGGCTTAAAGAAAGAGAGGTCTTATGCGAGACTCACATTTTTC 679
QY 661 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
DB 680 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 739
QY 721 TACTTAGTAAATAGCCTTCAGAGAAAGGCTCAGGAGATCATATTGTTGATGCTGTCAAA 780
DB 740 TACTTAGTAAATAGCCTTCAGAGAAAGGCTCAGGAGATCATATTGTTGATGCTGTCAAA 799
QY 781 ATTGAGCAGGAGTTTTTAAACAGAGGCTTTGCCAGTGGGCTCATTTGGAATGAATTCGATT 840
DB 800 ATTGAGCAGGAGTTTTTAAACAGAGGCTTTGCCAGTGGGCTCATTTGGAATGAATTCGATT 859
QY 841 TTGATGAAACAGTACATTTGAGTTTGTAGCTGACAGATTACTTGTGGACTTGGATTCTTCA 900
DB 860 TTGATGAAACAGTACATTTGAGTTTGTAGCTGACAGATTACTTGTGGACTTGGATTCTTCA 919
QY 901 AAGGTTTTTTCAGGAGAGAAATCCTTTTGTATTTTATGAAACATTTCTTTAGAGGAAAA 960
DB 920 AAGGTTTTTTCAGGAGAGAAATCCTTTTGTATTTTATGAAACATTTCTTTAGAGGAAAA 979
QY 961 ACAAATTTCTTTCAGAAACGAGTTTCAGAGTATCAGCGTTTTGCGAGTTATGCGAGAAACC 1020
DB 980 ACAAATTTCTTTCAGAAACGAGTTTCAGAGTATCAGCGTTTTGCGAGTTATGCGAGAAACC 1039
QY 1021 ACAGATAACGCTCTTCCACCTTGGATGACAGATTTT 1053
DB 1040 ACAGATAACGCTCTTCCACCTTGGATGACAGATTTT 1072
RESULT 5
BD064766 1081 bp DNA linear PAT 27-AUG-2002
LOCUS Novel protein and DNA thereof.
DEFINITION BD064766
ACCESSION BD064766
VERSION 1 GI:22610369
KEYWORDS JP 2001269184-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1081)
AUTHORS Nakamura,Y., Arakawa,H. and Tanaka,H.
TITLE Novel protein and DNA thereof
JOURNAL Patent: JP 2001269184-A 3 02-OCT-2001;
YUSUKE NAKAMURA,TAKEDA CHEMICAL INDUSTRIES LTD
COMMENT OS Homo sapiens (human)
PN JP 2001269184-A/3
PD 02-OCT-2001
PF 27-JUN-2000 JP 2000192401
PI YUSUKE NAKAMURA,HIROFUMI ARAKAWA,HIROSHI TANAKA PC
C12N15/09,A61K31/711,A61K38/00,A61K38/43,A61K45/00,A61K48/00, PC
A61P35/00,
PC A61P43/00,C07K16/32,C07K16/40,C12N1/15,C12N1/19,C12N1/21 PC
C12N5/10,C12N9/02,
PC C12P21/02,C12P21/08,G01N33/15,G01N33/50,G01N33/53//C12N15/09,
C12R1:91),
PC (C12N15/09,C12R1:19),(C12N9/02,C12R1:91),(C12P21/02,C12R1:91),
PC (C12P21/08,C12R1:91),C12N15/00,A61K37/02,A61K37/48,C12N5/00,
CC Novel protein and DNA thereof
FH Key Location/Qualifiers
FT source 1..1081
/organism="Homo sapiens"
/feature="genomic DNA"
/db_xref="taxon:9606"
FEATURES
source
1..1081
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 1053; DB 6; Length 1081;
Best Local Similarity 100.0%; Pred. No. 6.4e-255;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGCGACCGGAAAGCGGCGGCGGCTGGATCAGATCAGAGATCATCTTCA 60
DB 20 ATGGGCGACCGGAAAGCGGCGGCGGCTGGATCAGATCAGAGATCATCTTCA 79
QY 61 GACACCAACGAAAGTGAAATAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
DB 80 GACACCAACGAAAGTGAAATAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 139
QY 121 CGTTTGTGATCTTCCCAATCGATACCTGATATTTGGAATGATATTAACAGGACAG 180
DB 140 CGTTTGTGATCTTCCCAATCGATACCTGATATTTGGAATGATATTAACAGGACAG 199
QY 181 GCTTCTCTCTGACAGCAGAGAGTTGACTATCAAGAGATCTCTCACTGGAACAAG 240
DB 200 GCTTCTCTCTGACAGCAGAGAGTTGACTATCAAGAGATCTCTCACTGGAACAAG 259
QY 241 CTTAAGCAGATGAGAAGTACTTCTCTCACTCTTACATCTTACAGCTTTTTCGAGCAGTGAT 300
DB 260 CTTAAGCAGATGAGAAGTACTTCTCTCACTCTTACATCTTACAGCTTTTTCGAGCAGTGAT 319
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
DB 320 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 379
QY 361 CGCTGTTTCTATGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGATACAGTTTG 420
DB 380 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGATACAGTTTG 439
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 480
DB 440 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 499
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGATGAGATAGAGATAGAAAA 540
DB 500 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGATGAGATAGAGATAGAAAA 559
QY 541 TCTACTTTTGGGAAAGAGTGGGCTTTGCTGCTGATAGAGAGTTCCTCTCAGGA 600
DB 560 TCTACTTTTGGGAAAGAGTGGGCTTTGCTGCTGATAGAGAGTTCCTCTCAGGA 619
QY 601 TCTTTTGTGCTATATTCTGGCTTAAAGAAAGAGAGGTCTTATGCGAGACTCACATTTTTC 660
DB 620 TCTTTTGTGCTATATTCTGGCTTAAAGAAAGAGAGGTCTTATGCGAGACTCACATTTTTC 679
QY 661 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
DB 680 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 739
QY 721 TACTTAGTAAATAGCCTTCAGAGAAAGGCTCAGGAGATCATATTGTTGATGCTGTCAAA 780
DB 740 TACTTAGTAAATAGCCTTCAGAGAAAGGCTCAGGAGATCATATTGTTGATGCTGTCAAA 799
QY 781 ATTGAGCAGGAGTTTTTAAACAGAGGCTTTGCCAGTGGGCTCATTTGGAATGAATTCGATT 840
```

Qy	121	CGTTTGTTCATCTTTCCAAATCCAGTACCCTGTATATTGGAAATGTATAACAGGCACAG	180
Db	140	CGTTTGTTCATCTTTCCAAATCCAGTACCCTGTATATTGGAAATGTATAACAGGCACAG	199
Qy	181	GCTTCCTTCTGACAGAGAAGGTTGACTTATCAAGAGATCTCCCTCACTGGAACAAG	240
Db	200	GCTTCCTTCTGACAGAGAAGGTTGACTTATCAAGAGATCTCCCTCACTGGAACAAG	259
Qy	241	CTTAAAGCAGATGAGAAGTACTTCATCTCTCACATCTTAGCCCTTTTTCGAGCCAGTGAT	300
Db	260	CTTAAAGCAGATGAGAAGTACTTCATCTCTCACATCTTAGCCCTTTTTCGAGCCAGTGAT	319
Qy	301	GGAAATGTAAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT	360
Db	320	GGAAATGTAAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT	379
Qy	361	CGTGTTCCTATGCGCTTTTCAAAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG	420
Db	380	CGTGTTCCTATGCGCTTTTCAAAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG	439
Qy	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTGAA	480
Db	440	CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTGAA	499
Qy	481	ACCATGCCCTATGTTAAGAAAAGAGAGATTTGGCCCTTCGATGATAGCAGATAGAAA	540
Db	500	ACCATGCCCTATGTTAAGAAAAGAGAGATTTGGCCCTTCGATGATAGCAGATAGAAA	559
Qy	541	TCTACTTTTGGGAAAAGAGTGGTGGCCCTTCTGCTGTAGAGAGGATTTTCTTCTCAGGA	600
Db	560	TCTACTTTTGGGAAAAGAGTGGTGGCCCTTCTGCTGTAGAGAGGATTTTCTTCTCAGGA	619
Qy	601	AATGAACCTCATCAGCAGAGATGAAGACTTCACGTGACTTTTGCCTTGATGCTGTCCAA	660
Db	620	AATGAACCTCATCAGCAGAGATGAAGACTTCACGTGACTTTTGCCTTGATGCTGTCCAA	679
Qy	661	AATGAACCTCATCAGCAGAGATGAAGACTTCACGTGACTTTTGCCTTGATGCTGTCCAA	720
Db	680	AATGAACCTCATCAGCAGAGATGAAGACTTCACGTGACTTTTGCCTTGATGCTGTCCAA	739
Qy	721	TACTTAGTAAATAAGCCTTTCAGAAAGAAAGGCTCAGGAGATCATTTGATGCTGTCCAA	780
Db	740	TACTTAGTAAATAAGCCTTTCAGAAAGAAAGGCTCAGGAGATCATTTGATGCTGTCCAA	799
Qy	781	ATTGAGCAGAGTTTTAAACAGAGCCTTGCAGTTGGCTCATFTGGAATGAATTCAT	840
Db	800	ATTGAGCAGAGTTTTAAACAGAGCCTTGCAGTTGGCTCATFTGGAATGAATTCAT	859
Qy	841	TTGATCAACACGTACATTGAGTTTGTAGCTGACAGATTACTTGTGAACTTGGATTTCTCA	900
Db	860	TTGATCAACACGTACATTGAGTTTGTAGCTGACAGATTACTTGTGAACTTGGATTTCTCA	919
Qy	901	AAGGTTTTTTCAGCAGAAAATCCTTTTGATTTTATGGAATAATTTCTTTTAGAAGGAAA	960
Db	920	AAGGTTTTTTCAGCAGAAAATCCTTTTGATTTTATGGAATAATTTCTTTTAGAAGGAAA	979
Qy	961	ACAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTTATGCGAGAAACC	1020
Db	980	ACAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTTATGCGAGAAACC	1039
Qy	1021	ACAGATAAGCTTTCACCTTGGATGAGATTTT 1053	
Db	1040	ACAGATAAGCTTTCACCTTGGATGAGATTTT 1072	
RESULT 6			
BD093079			
LOCUS	BD093079	1081 bp	DNA linear PAT 27-AUG-2002
DEFINITION	Novel protein and its DNA.		
ACCESSION	BD093079		
VERSION	BD093079.1 GI:22638667		
KEYWORDS	WO 0100799-A/3.		
SOURCE	Homo sapiens (human)		

ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 1081)		
AUTHORS	Nakamura,Y., Arakawa,H. and Tanaka,H.		
TITLE	Novel protein and its DNA		
JOURNAL	Patent: WO 0100799-A 3 04-JAN-2001;		
COMMENT	TAKEDA CHEMICAL INDUSTRIES LTD, YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA		
OS	Homo sapiens (human)		
PN	WO 0100799-A/3		
PD	04-JAN-2001		
PF	27-JUN-2000 WO 2000JP004189		
PR	28-JUN-1999 JP 99P 181131, 06-JUL-1999 JP 99P 192391 PR		
PI	21-JAN-2000 JP 00P 017770		
PI	YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC		
C12N1/04, C12N15/53, C12N1/15, C12N1/19, C12N1/21, C12N5/10 PC			
, C12P21/02, A61K38/44,			
PC	A61K48/00, C07K16/40, G01N33/53, C12Q1/26		
CC			
PH	Key Location/Qualifiers.		
FEATURES	source		
ORIGIN	1.1081		
	/organism="Homo sapiens"		
	/mol type="genomic DNA"		
	/db_xref="taxon:9606"		
Query Match	100.0%; Score 1053; DB 6; Length 1081;		
Best Local Similarity	100.0%; Pred. No. 6.4e-255;		
Matches 1053; Conservative	0; Mismatches 0; Indels 0; Gaps 0;		
Qy	1	ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA	60
Db	20	ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA	79
Qy	61	GACACCAACCAAGTAGTAAATGAAATGAAAGAGCCACTCTTAAGAAAGAGTTCTCGC	120
Db	80	GACACCAACCAAGTAGTAAATGAAATGAAAGAGCCACTCTTAAGAAAGAGTTCTCGC	139
Qy	121	CGGTTTGTTCATCTTCCAAATCCAGTACCTGATATTTGGAAAATGTATAAACAGGCACAG	180
Db	140	CGGTTTGTTCATCTTCCAAATCCAGTACCTGATATTTGGAAAATGTATAAACAGGCACAG	199
Qy	181	GCTTCCTTCTGACAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAACAAG	240
Db	200	GCTTCCTTCTGACAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAACAAG	259
Qy	241	CTTAAAGCAGATGAGAAGTACTTCATCTCACATCTTAGCCCTTTTTCGAGCCAGTGAT	300
Db	260	CTTAAAGCAGATGAGAAGTACTTCATCTCACATCTTAGCCCTTTTTCGAGCCAGTGAT	319
Qy	301	GGAAATGTAAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT	360
Db	320	GGAAATGTAAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT	379
Qy	361	CGTGTTCCTATGCGCTTTTCAAAATTCATCAGAGATGTTCATCAGAGATGTACAGTTTG	420
Db	380	CGTGTTCCTATGCGCTTTTCAAAATTCATCAGAGATGTTCATCAGAGATGTACAGTTTG	439
Qy	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTGAA	480
Db	440	CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTGAA	499
Qy	481	ACCATGCCCTATGTTAAGAAAAGAGAGATTTGGCCCTTCGATGATAGCAGATAGAAA	540
Db	500	ACCATGCCCTATGTTAAGAAAAGAGAGATTTGGCCCTTCGATGATAGCAGATAGAAA	559
Qy	541	TCTACTTTTGGGAAAAGAGTGGTGGCCCTTCTGCTGTAGAGAGGATTTTCTTCTCAGGA	600
Db	560	TCTACTTTTGGGAAAAGAGTGGTGGCCCTTCTGCTGTAGAGAGGATTTTCTTCTCAGGA	619
Qy	601	TCTTTTGTCTCTATATTCTGGCTTAAAGAAAGAGAGGCTTTATGCGAGACTCACTTTTTC	660


```

COMMENT OS Homo sapiens (human)
PN JP 2001269184-A/1
PD 02-OCT-2001
PF 27-JUN-2000 JP 2000192401
PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
C12N15/09, A61K31/11, A61K38/00, A61K38/43, A61K45/00, A61K48/00, PC
A61P35/00,
PC A61P43/00, C07K16/32, C07K16/40, C12N1/15, C12N1/19, C12N1/21 PC
, C12N5/10, C12N9/02,
PC C12P21/02, C12P21/08, G01N33/15, G01N33/50, G01N33/53// (C12N15/09,
C12R1/91),
PC (C12N15/09, C12R1/19), (C12N9/02, C12R1/91), (C12P21/02, C12R1/91),
PC (C12P21/08, C12R1/91), (C12N15/00, A61K37/02, A61K37/48, C12N5/00,
C12N15/00, C12R1/91), (C12N15/00, C12R1/19)
CC Novel protein and DNA thereof
FH Key Location/Qualifiers
FT source 1..1053
FT /organism='Homo sapiens (human)'.
FEATURES
source
1..1053
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.8%; Score 1051.4; DB 6; Length 1053;
Best Local Similarity 99.9%; Pred. No. 1.6e-254;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGCGACCCCGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
DB 1 ATGGCGACCCCGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
QY 61 GACACCAAGCAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAGAGTTCTCGC 120
DB 61 GACACCAAGCAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAGAGTTCTCGC 120
QY 121 CGGTTTGTCATCTTCCATCCAGTACCTGATATTTGGAATGATTAACAGGCACAG 180
DB 121 CGGTTTGTCATCTTCCATCCAGTACCTGATATTTGGAATGATTAACAGGCACAG 180
QY 181 GCTTCCTCTGACAGCAGAGAGTTGACTTATCAAGATCTCCCTCACTGGAACAAG 240
DB 181 GCTTCCTCTGACAGCAGAGAGTTGACTTATCAAGATCTCCCTCACTGGAACAAG 240
QY 241 CTTAAGCAGATGAGAAGTACTTCACTCTCAATCTTAGCCCTTTTTCGAGCCAGTGAT 300
DB 241 CTTAAGCAGATGAGAAGTACTTCACTCTCAATCTTAGCCCTTTTTCGAGCCAGTGAT 300
QY 301 GGAATGTGAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
DB 301 GGAATGTGAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
QY 361 CGCTGTTCTATGGCTTTCAAAATTCATCGAGAATGTTCACTCAGAGATGACAGTTTG 420
DB 361 CGCTGTTCTATGGCTTTCAAAATTCATCGAGAATGTTCACTCAGAGATGACAGTTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
DB 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATGTTGGGCTTCGATGGATGAGATAGAAAA 540
DB 481 ACCATGCCCTATGTTAAGAAAAAGCAGATGTTGGGCTTCGATGGATGAGATAGAAAA 540
QY 541 TCTACTTTTGGGAAAGAGTGTGGCTTTTGTGCTGTAGAAGGAGTTTCTTCTCAGGA 600
DB 541 TCTACTTTTGGGAAAGAGTGTGGCTTTTGTGCTGTAGAAGGAGTTTCTTCTCAGGA 600
QY 601 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTCC 660
DB 601 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTCC 660

```

```

QY 661 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
DB 661 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
QY 721 TACTTAGTAATAAGCCCTTCAGAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 780
DB 721 TACTTAGTAATAAGCCCTTCAGAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 780
QY 781 ATTGAGCAGAGGTTTTTAAACAGAGCCCTTGGCCAGTTGGCCCTCATTTGGAATGAATTGCATT 840
DB 781 ATTGAGCAGAGGTTTTTAAACAGAGCCCTTGGCCAGTTGGCCCTCATTTGGAATGAATTGCATT 840
QY 841 TTGATGAACACAGTACATTTAGTCTTGTAGCTGACAGATTTACTTGTGAACTTTGGATTTCTCA 900
DB 841 TTGATGAACACAGTACATTTAGTCTTGTAGCTGACAGATTTACTTGTGAACTTTGGATTTCTCA 900
QY 901 AAGTTTTTTCAGGAGAAAAATCCTTTTGAATTTTATGAAAAACATTTCTTTAGAGAAAA 960
DB 901 AAGTTTTTTCAGGAGAAAAATCCTTTTGAATTTTATGAAAAACATTTCTTTAGAGAAAA 960
QY 961 ACAAAATTTCTTTCAGAAACGAGTTTCAGAGTATCAGCGTTTTCGAGTTATGCGAGAAACC 1020
DB 961 ACAAAATTTCTTTCAGAAACGAGTTTCAGAGTATCAGCGTTTTCGAGTTATGCGAGAAACC 1020
QY 1021 ACAGATAACGCTTTCACCTTTGGATGAGATTTT 1053
DB 1021 ACAGATAACGCTTTCACCTTTGGATGAGATTTT 1053
RESULT 9
BD093077
LOCUS BD093077
DEFINITION Novel protein and its DNA.
ACCESSION BD093077
VERSION BD093077.1 GI:22638665
KEYWORDS WO 0100799-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1053)
AUTHORS Nakamura, Y., Arakawa, H. and Tanaka, H.
TITLE Novel protein and its DNA
JOURNAL Patent: WO 0100799-A 1 04-JAN-2001;
TAKEDA CHEMICAL INDUSTRIES LTD, YUSUKE NAKAMURA, HIROFUMI ARAKAWA,
HIROSHI TANAKA
COMMENT OS Homo sapiens (human)
PN WO 0100799-A/1
PD 04-JAN-2001
PF 27-JUN-2000 WO 2000JP004189
PR 28-JUN-1999 JP 99P 181131, 06-JUL-1999 JP 99P 192391 PR
21-JAN-2000 JP 00P 017770
PI YUSUKE NAKAMURA, HIROFUMI ARAKAWA, HIROSHI TANAKA PC
C12N9/04, C12N15/53, C12N1/15, C12N1/19, C12N1/21, C12N5/10 PC
, C12P21/02, A61K38/44,
PC A61K48/00, C07K16/40, G01N33/53, C12Q1/26
CC
FH Key Location/Qualifiers.
FEATURES
source
1..1053
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.8%; Score 1051.4; DB 6; Length 1053;
Best Local Similarity 99.9%; Pred. No. 1.6e-254;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGCGACCCCGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
DB 1 ATGGCGACCCCGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60

```


61 GACACAAAGAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db GACACAAAGAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
QY CGGTTTGTCTATCTTCCATCCAGTACCTCTGATATTTGGAAATGTATAAACAGGCACAG 180
Db CGGTTTGTCTATCTTCCATCCAGTACCTCTGATATTTGGAAATGTATAAACAGGCACAG 180
QY GCTTCTCTCTGGACAGAGAGGTTGACTTATCAAAAGATCTCCCTCACTGGAACAAG 240
Db GCTTCTCTCTGGACAGAGAGGTTGACTTATCAAAAGATCTCCCTCACTGGAACAAG 240
QY CTTAAAGCAGATGAGAAGTACTCTCTCAATCTTAGCCCTTTTTCAGCCAGTGAT 300
Db CTTAAAGCAGATGAGAAGTACTCTCTCAATCTTAGCCCTTTTTCAGCCAGTGAT 300
QY GGAATTTGAATGAATAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCCAGAGGCT 360
Db GGAATTTGAATGAATAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCCAGAGGCT 360
QY CGCTGTTTCTATGCTTTCAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db CGCTGTTTCTATGCTTTCAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
QY CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAAATGCAATTGAA 480
Db CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAAATGCAATTGAA 480
QY ACATGCTCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGGATGGATAGCAGATAGAAA 540
Db ACATGCTCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGGATGGATAGCAGATAGAAA 540
QY TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 600
Db TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 600
QY TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 660
Db TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 660
QY AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
QY TACTTAGTAAATGAAGCTTTCAGAAAGAGGTTCAGGAGATCATTTGTTGATGCTGTCAA 780
Db TACTTAGTAAATGAAGCTTTCAGAAAGAGGTTCAGGAGATCATTTGTTGATGCTGTCAA 780
QY ATTGAGCAGGAGTTTAAACAGAGCTTGCAGAGCTTGGCTCATTTGGAATGGAATTCATT 840
Db ATTGAGCAGGAGTTTAAACAGAGCTTGCAGAGCTTGGCTCATTTGGAATGGAATTCATT 840
QY TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATTGTGGAACCTTGGATTTCA 900
Db TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATTGTGGAACCTTGGATTTCA 900
QY AAGGTTTTTCAGCAGAAAAATCTTTTGAATTTATGGAAGAACATTTCTTTAGAGAAAA 960
Db AAGGTTTTTCAGCAGAAAAATCTTTTGAATTTATGGAAGAACATTTCTTTAGAGAAAA 960
QY ACAAAATTTCTTTGAGAAACAGTTTTCAGAGTATCAGCGTTTTCAGGATGATGAGAAACC 1020
Db ACAAAATTTCTTTGAGAAACAGTTTTCAGAGTATCAGCGTTTTCAGGATGATGAGAAACC 1020
QY ACAGATAACGTTCTTCACTTGGATGAGATTTT 1053
Db ACAGATAACGTTCTTCACTTGGATGAGATTTT 1053

RESULT 10
CQ714252
LOCUS
DEFINITION Sequence 186 from Patent WO02068579.
CQ714252 1056 bp DNA linear PAT 03-FEB-2004

ACCESSION CQ714252
VERSION CQ714252.1 GI:42275109
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of human exons or transcripts, for detecting expression and other uses thereof
JOURNAL Patent: WO 02068579-A 186 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
source
1..1056
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.8%; Score 1051.4; DB 6; Length 1056;
Best Local Similarity 99.9%; Pred. No. 1.6e-254;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGGCGACCCGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 ATGGGCGACCCGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
QY 61 GACACAAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 61 GACACAAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
QY 121 CGGTTTGTCTATCTTCCAAATCCAGTACCTGATATTTGGAAATGTATAACAGGCACAG 180
Db 121 CGGTTTGTCTATCTTCCAAATCCAGTACCTGATATTTGGAAATGTATAACAGGCACAG 180
QY 181 GCTTCTCTTGGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 240
Db 181 GCTTCTCTTGGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 240
QY 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCAGATCTTAGCCCTTTTTCAGCCAGTGAT 300
Db 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCAGATCTTAGCCCTTTTTCAGCCAGTGAT 300
QY 301 GGAATTTGAATGAATAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCCAGAGGCT 360
Db 301 GGAATTTGAATGAATAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCCAGAGGCT 360
QY 361 CGCTGTTTCTATGCTTTCAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGCTTTCAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAAATGCAATTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAAATGCAATTGAA 480
QY 481 ACCATGCCCTATCTTAAGAAAAAGCAGATTTGGCCCTTGCATGGAATGAGATAGAAAA 540
Db 481 ACCATGCCCTATCTTAAGAAAAAGCAGATTTGGCCCTTGCATGGAATGAGATAGAAAA 540
QY 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 600
QY 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 660
Db 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 660
QY 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 660
Db 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 660
QY 661 AATGAACCTCATCAGCAGAGATGAAGAGCTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 661 AATGAACCTCATCAGCAGAGATGAAGAGCTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720


```

QY 721 TACTTAGTAAATAAGCCTTCAGAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAA 780
DB 721 TACTTAGTAAATAAGCCTTCAGAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAA 780
QY 781 ATTGACGAGGCTTTTAAACAGAGCCTTGCAGTTGGCTCATTTGGAATGAATTGCATT 840
DB 781 ATTGACGAGGCTTTTAAACAGAGCCTTGCAGTTGGCTCATTTGGAATGAATTGCATT 840
QY 841 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
DB 841 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
QY 901 AAGGTTTTTTCAGCAGAAAATCCTTTTGTATTTATGGAACAATCTTTTGTAGAGGAAA 960
DB 901 AAGGTTTTTTCAGCAGAAAATCCTTTTGTATTTATGGAACAATCTTTTGTAGAGGAAA 960
QY 961 ACAAAATTTCTTGTAGAAAACGAGTTTCAGAGTATCAGCGTTTGTGAGTATGCGAGAAAC 1020
DB 961 ACAAAATTTCTTGTAGAAAACGAGTTTCAGAGTATCAGCGTTTGTGAGTATGCGAGAAAC 1020
QY 1021 ACAGATAACGCTTTCACCTTGGATGACAGATTTT 1053
DB 1021 ACAGATAACGCTTTCACCTTGGATGACAGATTTT 1053

RESULT 11
AB166671
LOCUS AB166671 1056 bp mRNA linear PRI 09-MAR-2004
DEFINITION Homo sapiens p53R2 mRNA for p53-inducible ribonucleotide reductase
small subunit 2, complete cds.
ACCESSION AB166671
VERSION AB166671.1 GI:45259568
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Ugai,H. and Yokoyama,K.K.
TITLE Homo sapiens p53R2 mRNA for p53-inducible ribonucleotide reductase
small subunit 2, complete cds
JOURNAL Published Only in Database (2004)
REFERENCE 2
AUTHORS Ugai,H. and Yokoyama,K.K.
TITLE Direct Submission
JOURNAL Submitted (05-MAR-2004) Hideyo Ugai, RIKEN BioResource Center, Gene
Engineering Division; 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074,
Japan [E-mail:ugai@rtc.riken.go.jp, Tel:81-29-836-3612,
Fax:81-29-836-9120]
FEATURES
source
1..1056
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="p53R2.C19"
/cell_line="Saos-2"
/location="Osteogenic sarcoma"
1..1056
/gene="p53R2"
1..1056
/gene="p53R2"
/codon_start=1
/product="p53-inducible ribonucleotide reductase small
subunit 2"
/protein_id="BAD12267.1"
/db_xref="GI:45259569"
/translation="MGDPERPEAAGLDQDERSSTNSESEIKNSEPELLKSRPFVI
FPIQPIWKMTQQAQSFWTAEEVDLSKDLPHWNKLKADKFIYFISHILAFPAASDGI
VNEMLVFSEVQPEARCFYQFQILNENHSEMYSLIDITYIRDPKKREFLPAIE
TMPVYKKADWALRIADKSTFGERVVAFVAFVFFSGSPAAIFWLKKRGLMPEGLT
FSNELISDEGLHCDIFACLMFOYLVNKPSEERVRIIVDAVKIIOEFLFEALPVGLIG
MNCILMKOYIEFVADRLVLELGFSGVQENPFDPFENISLBGNTNFFEKRVSEVQRF
AVMAETTONVFLDADF"

```

ORIGIN

```

Query Match 99.8%; Score 1051.4; DB 9; Length 1056;
Best Local Similarity 99.9%; Pred. No. 1.6e-254;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGCAGCCGGAAGCGCGGAGCGCGGCTGAGTATGAGATGAGAGATCATCTTCA 60
DB 1 ATGGCGCAGCCGGAAGCGCGGAGCGCGGCTGAGTATGAGATGAGAGATCATCTTCA 60
QY 61 GACACCAACGAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
DB 61 GACACCAACGAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
QY 121 CGGTTGTCTATCTTTCCATCCAGTACCTGATATTTGGAAAATGTATAAAGAGGACAG 180
DB 121 CGGTTGTCTATCTTTCCATCCAGTACCTGATATTTGGAAAATGTATAAAGAGGACAG 180
QY 181 GCTTTCCTTCTGGACAGAGAGGTTGACTTATCAAAAGGATCTCCCTCACTGGAAACAAG 240
DB 181 GCTTTCCTTCTGGACAGAGAGGTTGACTTATCAAAAGGATCTCCCTCACTGGAAACAAG 240
QY 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCAATCTTTAGCCCTTTTTCAGCCAGTGTAT 300
DB 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCAATCTTTAGCCCTTTTTCAGCCAGTGTAT 300
QY 301 GGAATTGTAATGAANAATTTGGTGGAGGCTTTTAGTCAGGAGGTCAGGTTCCAGAGGCT 360
DB 301 GGAATTGTAATGAANAATTTGGTGGAGGCTTTTAGTCAGGAGGTCAGGTTCCAGAGGCT 360
QY 361 CGCTGTTTCTATGCTTTTCAAAATTTCTCATCAGAAATGTTTCACTCAGAGATGACAGTTTG 420
DB 361 CGCTGTTTCTATGCTTTTCAAAATTTCTCATCAGAAATGTTTCACTCAGAGATGACAGTTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATCAATTTGAA 480
DB 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATCAATTTGAA 480
QY 481 ACCATGCCCTATGTTTAAAGAAAAGAGATTTGGGCTTTCGATGATGATGAGATGAGAAA 540
DB 481 ACCATGCCCTATGTTTAAAGAAAAGAGATTTGGGCTTTCGATGATGATGAGATGAGAAA 540
QY 541 TCTACTTTTGGGGAAGAGTGTGGCTTCTCTCTGTAGAGGAGTTTCTTCTCAGGA 600
DB 541 TCTACTTTTGGGGAAGAGTGTGGCTTCTCTCTGTAGAGGAGTTTCTTCTCAGGA 600
QY 601 TCTTTTGTCTCTATATTTCTGGCTAAAGAGAGGTCCTTATGCCAGGAGTCTTATGCCAGGACT 660
DB 601 TCTTTTGTCTCTATATTTCTGGCTAAAGAGAGGTCCTTATGCCAGGAGTCTTATGCCAGGACT 660
QY 661 AATGAATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGTGCTGCTGATGTTCCAA 720
DB 661 AATGAATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGTGCTGCTGATGTTCCAA 720
QY 721 TACTTAGTAAATAAGCCTTCAGAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAA 780
DB 721 TACTTAGTAAATAAGCCTTCAGAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAA 780
QY 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTGTGGCCCTCATTTGGAATGAATTGCATT 840
DB 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTGTGGCCCTCATTTGGAATGAATTGCATT 840
QY 841 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
DB 841 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
QY 901 AAGGTTTTTTCAGCAGAAAATCCTTTTGTATTTATGGAACAATCTTTTGTAGAGGAAA 960
DB 901 AAGGTTTTTTCAGCAGAAAATCCTTTTGTATTTATGGAACAATCTTTTGTAGAGGAAA 960
QY 961 ACAAAATTTCTTGTAGAAAACGAGTTTCAGAGTATCAGCGTTTGTGAGTATGCGAGAAAC 1020
DB 961 ACAAAATTTCTTGTAGAAAACGAGTTTCAGAGTATCAGCGTTTGTGAGTATGCGAGAAAC 1020

```


Best Local Similarity 99.9%; Pred. No. 1.5e-254;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGAAAGCGCGAAGCGCGGCTGGATCAGGATCATCTTCA 60
DB 245 ATGGGCGACCCGAAAGCGCGAAGCGCGGCTGGATCAGGATCATCTTCA 304

QY 61 GACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
DB 305 GACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 364

QY 121 CGTTTGTCTATCTTCCAAATCCAGTACCTGATATTTGGAAATGATATAACAGGCACAG 180
DB 365 CGTTTGTCTATCTTCCAAATCCAGTACCTGATATTTGGAAATGATATAACAGGCACAG 424

QY 181 GCTTCTCTTGGACAGAGAGGTTGACTTATCAAGAGTCTCCCTCACTGGACAG 240
DB 425 GCTTCTCTTGGACAGAGAGGTTGACTTATCAAGAGTCTCCCTCACTGGACAG 484

QY 241 CTTTAAAGCAGATCAGAGAGTACTTCTCATCATCTTTAGCCTTTTTCAGCCAGTGTAT 300
DB 485 CTTTAAAGCAGATCAGAGAGTACTTCTCATCATCTTTAGCCTTTTTCAGCCAGTGTAT 544

QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTTAGTCAGGAGGTGCAGGCT 360
DB 545 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTTAGTCAGGAGGTGCAGGCT 604

QY 361 CGCTGTTTCTATGGCTTTCAAAATTCATCGAGAAATGTTTCACTCAGAGATGTACGTTTG 420
DB 605 CGCTGTTTCTATGGCTTTCAAAATTCATCGAGAAATGTTTCACTCAGAGATGTACGTTTG 664

QY 421 CTGATAGACATTTACATCAGATCCCAAGAAAGGAAATTTTATTAATGCAATTGAA 480
DB 665 CTGATAGACATTTACATCAGATCCCAAGAAAGGAAATTTTATTAATGCAATTGAA 724

QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGGATGGATAGCAGATAGAAA 540
DB 725 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGGATGGATAGCAGATAGAAA 784

QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGAAGAGAGTTTCTCTCAGGA 600
DB 785 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGAAGAGAGTTTCTCTCAGGA 844

QY 601 TCTTTTGTCTATATCTTGGCTTAAGAAAGAGAGTCTTATGCGCAGACTCTCTTTTCC 660
DB 845 TCTTTTGTCTATATCTTGGCTTAAGAAAGAGAGTCTTATGCGCAGACTCTCTTTTCC 904

QY 661 AATGAATCTCATCAGCAGAGATGAAGACTTCTACTGTGACTTTTCTGCTGCTGATGTTCCAA 720
DB 905 AATGAATCTCATCAGCAGAGATGAAGACTTCTACTGTGACTTTTCTGCTGCTGATGTTCCAA 964

QY 721 TACTTAGTAAATTAAGCTTTCAGAAAGAGGTCAGGAGATCATTTGTGATGCTGTCAA 780
DB 965 TACTTAGTAAATTAAGCTTTCAGAAAGAGGTCAGGAGATCATTTGTGATGCTGTCAA 1024

QY 781 ATTGAGCAGAGTTTATTAAGAGCCTTGGCAGTGGCTCATTTGGAATGGAATTCAT 840
DB 1025 ATTGAGCAGAGTTTATTAAGAGCCTTGGCAGTGGCTCATTTGGAATGGAATTCAT 1084

QY 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTTGTGAACTTGGATTTCTCA 900
DB 1085 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTTGTGAACTTGGATTTCTCA 1144

QY 901 AAGGTTTTTTCAGGCAGAAATCTTTTGAATTTTATGGAACAACTTTCTTTAGAGGAA 960
DB 1145 AAGGTTTTTTCAGGCAGAAATCTTTTGAATTTTATGGAACAACTTTCTTTAGAGGAA 1204

QY 961 ACAATTTCTTTGAGAAACAGTTTTCAGAGTATCAGGTTTTCAGGTTTATGCGAGAAACC 1020
DB 1205 ACAATTTCTTTGAGAAACAGTTTTCAGAGTATCAGGTTTTCAGGTTTATGCGAGAAACC 1264

QY 1021 ACAGATAACGCTTTCACCTTGGATCAGATTTT 1053

Db 1265 ACAGATAACGCTTTCACCTTGGATCAGATTTT 1297

RESULT 14
BD093078
LOCUS
DEFINITION
Novel protein and its DNA.
ACCESSION
BD093078
VERSION
BD093078.1 GI:22638666
KEYWORDS
WO 0100799-A/2;
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
1 (bases 1 to 4955)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS
Nakamura,Y., Arakawa,H. and Tanaka,H.
TITLE
Novel protein and its DNA
JOURNAL
TAKEDA CHEMICAL INDUSTRIES LTD, YUSUKE NAKAMURA, HIROFUMI ARAKAWA,
HIROSHI TANAKA
COMMENT
OS Homo sapiens (human)
PN WO 0100799-A/2
PD 04-JAN-2001
PF 27-JUN-2000 WO 2000JP004189
PR 28-JUN-1999 JP 99P 181131.06-JUL-1999 JP 99P 192391 PR
21-JAN-2000 JP 00P 017770
PI YUSUKE NAKAMURA,HIROFUMI ARAKAWA,HIROSHI TANAKA PC
C12N9/04,C12N15/53,C12N1/15,C12N1/19,C12N1/21,C12N5/10 PC
C12P21/02,A61K38/44,
PC A61K48/00,C07K16/40,G01N33/53,C12Q1/26
CC

FEATURES
source
FH Key Location/Qualifiers.
1.4955
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 99.8%; Score 1051.4; DB 6; Length 4955;
Best Local Similarity 99.9%; Pred. No. 1.5e-254;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGAAAGCGCGAAGCGCGGCTGGATCAGGATCATCTTCA 60
DB 245 ATGGGCGACCCGAAAGCGCGAAGCGCGGCTGGATCAGGATCATCTTCA 304

QY 61 GACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
DB 305 GACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 364

QY 121 CGTTTGTCTATCTTCCAAATCCAGTACCTGATATTTGGAAATGATATAACAGGCACAG 180
DB 365 CGTTTGTCTATCTTCCAAATCCAGTACCTGATATTTGGAAATGATATAACAGGCACAG 424

QY 181 GCTTCTCTTGGACAGAGAGGTTGACTTATCAAGAGTCTCCCTCACTGGACAG 240
DB 425 GCTTCTCTTGGACAGAGAGGTTGACTTATCAAGAGTCTCCCTCACTGGACAG 484

QY 241 CTTTAAAGCAGATCAGAGAGTACTTCTCATCATCTTTAGCCTTTTTCAGCCAGTGTAT 300
DB 485 CTTTAAAGCAGATCAGAGAGTACTTCTCATCATCTTTAGCCTTTTTCAGCCAGTGTAT 544

QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTTAGTCAGGAGGTGCAGGCT 360
DB 545 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTTAGTCAGGAGGTGCAGGCT 604

QY 361 CGCTGTTTCTATGGCTTTCAAAATTCATCGAGAAATGTTTCACTCAGAGATGTACGTTTG 420
DB 605 CGCTGTTTCTATGGCTTTCAAAATTCATCGAGAAATGTTTCACTCAGAGATGTACGTTTG 664

QY 421 CTGATAGACATTTACATCAGATCCCAAGAAAGGAAATTTTATTAATGCAATTGAA 480

665 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA 724
QY 481 ACCATGCCCTATGTTAAGAAAACAGAGATGGGCTTCGGATGGATAGATAGAAAA 540
Db 725 ACCATGCCCTATGTTAAGAAAACAGAGATGGGCTTCGGATGGATAGATAGAAAA 784
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAGAGGATTTTCTTCAGGA 600
Db 795 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAGAGGATTTTCTTCAGGA 844
QY 601 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGTCTTATGCGAGATCTCTTTTCC 660
Db 845 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGTCTTATGCGAGATCTCTTTTCC 904
QY 661 ATGAACTCATCAGCAGAGATGAGACATCTACTGTGACTTTGCTGCTGCTGATGTTCCAA 720
Db 905 AATGAACATCATCAGCAGAGATGAGACATCTACTGTGACTTTGCTGCTGCTGATGTTCCAA 964
QY 721 TACTTAGTAAATAAGCCTTCAGAGAAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 780
Db 965 TACTTAGTAAATAAGCCTTCAGAGAAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 1024
QY 781 ATGAGCAGGAGTCTTTTAAAGAGCCTTGCAGTGGCTCATTTGGAATGAATTCATT 840
Db 1025 ATTGAGCAGGAGTCTTTTAAAGAGCCTTGCAGTGGCTCATTTGGAATGAATTCATT 1084
QY 841 TTGATGAAACAGTACATTCAGTCTGAGAGATTAATCTTGTGAACTTGGATTTCTCA 900
Db 1085 TTGATGAAACAGTACATTCAGTCTGAGAGATTAATCTTGTGAACTTGGATTTCTCA 1144
QY 901 AAGGTTTTTCAGCAGAGAAATCCTTTGATTTTATGGAATAATTTCTTTAGAGAAAA 960
Db 1145 AAGGTTTTTCAGCAGAGAAATCCTTTGATTTTATGGAATAATTTCTTTAGAGAAAA 1204
QY 961 ACAATTTCTTTGAGAAACAGTCTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC 1020
Db 1205 ACAATTTCTTTGAGAAACAGTCTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC 1264
QY 1021 ACAGATAAGCTTTCACCTTGATGAGATTTT 1053
Db 1265 ACAGATAAGCTTTCACCTTGATGAGATTTT 1297

RESULT 15
AB036063
LOCUS Homo sapiens p53R2 mRNA 4955 bp mRNA linear PRI 11-MAR-2000
DEFINITION Homo sapiens p53R2 mRNA for ribonucleotide reductase, complete cds.
ACCESSION AB036063
VERSION AB036063.1 GI:7229085
KEYWORDS p53R2; ribonucleotide reductase.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (sites)
AUTHORS Tanaka.H., Arakawa.H., Yamaguchi.T., Shiraishi.K., Fukuda.S., Matsui.K., Takei.Y. and Nakamura.Y.
TITLE A ribonucleotide reductase gene involved in a p53-dependent cell-cycle checkpoint for DNA damage
JOURNAL Nature 404 (6773), 42-49 (2000)
MEDLINE 20179179
PUBMED 10716435
REFERENCE 2 (bases 1 to 4955)
AUTHORS Tanaka.H., Arakawa.H. and Nakamura.Y.
TITLE Direct Submission
JOURNAL Submitted (18-DEC-1999) Yusuke Nakamura, University of Tokyo, Institute of Medical Science, Human Genome Center, Laboratory of Molecular Medicine; 4-6-1 Shiokanadai, Minato-ku, Tokyo 108-8639, Japan [E-mail:yusuke@ims.u-tokyo.ac.jp, Tel:+81-3-5449-5372, Fax:+81-3-5449-5433]
FEATURES
1. 4955
/organism="Homo sapiens"
source

/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="8"
/map="8q23.1"
1. 4955
/gene="p53R2"
245. 1300
/gene="p53R2"
/codon_start=1
/product="ribonucleotide reductase"
/protein_id="BAA92434.1"
/db_xref="GI:7229086"
/translation="MGDPPEPPEAAGLDQDERSSDTNESEIKSNEEPLKRSRRFVI
FPIQYDPIWKMYKQAOAFTWABEVDLSKDLPHNKLKADKXYFISHILAFPAASDGI
VNEVLVERSEQVQVPEARCFYGOILLIENVHSEMYSLIDTYIRDPKRFLFNAIE
THPYVKKADWALRWIADRSKSTGERVVAFAAEGVEFSSFAIPLWKKRGLMPLGT
FNEELISRDGELHCDFAFLMFQVLNPKPSBERVEIIVDAVKIIOEFUTEALPVLGI
MNCILMKQYIEFVADRLVELGFSKVQAEENPFDMENI SLEKTNFFEKRVSEYQRF
AVMAETTDNVFTLDADF"
4933
/gene="p53R2"
/note="22 a nucleotides"
polyA_site
ORIGIN
Query Match 99.8%; Score 1051.4; DB 9; Length 4955;
Best Local Similarity 99.9%; Pred. No. 1.5e-254;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGCGACCCCGAAAGGCGGAGCGCGGCTGCATCAGGATGAGAGATCATCTTCA 60
Db 245 ATGGCGACCCCGAAAGGCGGAGCGCGGCTGCATCAGGATGAGAGATCATCTTCA 304
QY 61 GACACCAACGAAAGTGAATATAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTCTCGC 120
Db 305 GACACCAACGAAAGTGAATATAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTCTCGC 364
QY 121 CGGTTGTTCATCTTTCCAAATCCAGTACCCCTGATATTTGGAAAATGTATAAACAGGCACAG 180
Db 365 CGGTTGTTCATCTTTCCAAATCCAGTACCCCTGATATTTGGAAAATGTATAAACAGGCACAG 424
QY 181 GCTTCTCTTTCGACAGCAGAGAGTGTGACTTATCAAAAGAGTCTCCCTCACTGGAACAAG 240
Db 425 GCTTCTCTTTCGACAGCAGAGAGTGTGACTTATCAAAAGAGTCTCCCTCACTGGAACAAG 484
QY 241 CTTAAAGCAGATCAGAGATGACTTCACTCTCACTCTAGCCCTTTTTCGAGCCAGTGTAT 300
Db 485 CTTAAAGCAGATCAGAGATGACTTCACTCTCACTCTAGCCCTTTTTCGAGCCAGTGTAT 544
QY 301 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGAGTGCAGGTTCCAGAGGCT 360
Db 545 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGAGTGCAGGTTCCAGAGGCT 604
QY 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCAGAAATGTTCATCAGAGATGTACAGTTTG 420
Db 605 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCAGAAATGTTCATCAGAGATGTACAGTTTG 664
QY 421 CTGATAGACATTTACATCAGATCCCAAGAAAAGGAATTTTATTAATGCAATTGAA 480
Db 665 CTGATAGACATTTACATCAGATCCCAAGAAAAGGAATTTTATTAATGCAATTGAA 724
QY 481 ACCATGCCCTATGTTAAGAAAACAGAGATTTGGGCTTTCGATGATAGATAGAAAA 540
Db 725 ACCATGCCCTATGTTAAGAAAACAGAGATTTGGGCTTTCGATGATAGATAGAAAA 784
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAAAGAGTCTTCTTCAGGA 600
Db 785 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAAAGAGTCTTCTTCAGGA 844
QY 601 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGTCTTATGCGAGGACTCACTTTTCC 660
Db 845 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGTCTTATGCGAGGACTCACTTTTCC 904
QY 661 AATGAACATCATCAGCAGAGATGAGAGTCTTCACTGTGACTTTGCTGCTGATGTTCCAA 720

```
Db      |||||||
905 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 964
Qy      |||||||
721 TACTTAGTAAATAAGCCTTCAGAAGAAAGGTCAGGAGATCATTTGATGCTGTCAA 780
Db      |||||||
965 TACTTAGTAAATAAGCCTTCAGAAGAAAGGTCAGGAGATCATTTGATGCTGTCAA 1024
Qy      |||||||
781 ATTGAGCAGGAGTTTAAACAGAAGCCTTGCCAGTTGGCCTCATTTGGAATGAATTGCATT 840
Db      |||||||
1025 ATTGAGCAGGAGTTTAAACAGAAGCCTTGCCAGTTGGCCTCATTTGGAATGAATTGCATT 1084
Qy      |||||||
841 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTTGTGAACTTGGATTCTCA 900
Db      |||||||
1085 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTTGTGAACTTGGATTCTCA 1144
Qy      |||||||
901 AAGGTTTTTCAGGCAGAAAATCCTTTTGATTTTATGGAACAATTTCTTTAGAGGAAAA 960
Db      |||||||
1145 AAGGTTTTTCAGGCAGAAAATCCTTTTGATTTTATGGAACAATTTCTTTAGAGGAAAA 1204
Qy      |||||||
961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020
Db      |||||||
1205 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1264
Qy      |||||||
1021 ACAGATAACGTTCTTCACTTGGATGAGATTTT 1053
Db      |||||||
1265 ACAGATAACGTTCTTCACTTGGATGAGATTTT 1297
```

Search completed: October 30, 2005, 01:51:31
Job time : 3261 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2005, 21:38:30 ; Search time 440.5 Seconds
(without alignments)
14150.937 Million cell updates/sec

Title: US-10-698-228-12

Perfect score: 1053

Sequence: 1 atggcgaccgcgaaagcc.....tcacctggatgcagatttt 1053

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: _geneseqn1980s:*
- 2: _geneseqn1990s:*
- 3: _geneseqn2000s:*
- 4: _geneseqn2001as:*
- 5: _geneseqn2001bs:*
- 6: _geneseqn2002as:*
- 7: _geneseqn2002bs:*
- 8: _geneseqn2003as:*
- 9: _geneseqn2003bs:*
- 10: _geneseqn2003cs:*
- 11: _geneseqn2003ds:*
- 12: _geneseqn2004as:*
- 13: _geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1053	100.0	1053	4 AAF32447	Aaf32447 Human rib
2	1053	100.0	1081	4 AAF32440	Aaf32440 Human rib
3	1051.4	99.8	1053	4 AAF32438	Aaf32438 Human rib
4	1051.4	99.8	2596	3 AAI2411	Aai2411 cDNA enco
5	1051.4	99.8	4955	4 AAF32439	Aaf32439 Human rib
6	1051.4	99.8	4955	13 ADR24210	Adr24210 Breast ca
7	1048.2	99.5	1601	4 AAI14924	Aai14924 Human cDN
8	650.6	61.8	706	13 ADQ57092	Adq57092 Novel can
9	580	55.1	2482	4 AAS44917	Aas44917 Human con
10	580	55.1	2500	6 ABL65414	Abi65414 Lung canc
11	580	55.1	2500	6 ABL66517	Abi66517 Lung canc
12	580	55.1	2500	6 ABL65859	Abi65859 Lung canc
13	580	55.1	2500	8 ABX10335	Abx10335 DNA enco
14	580	55.1	2500	11 ADI32044	Adi32044 Human cDN
15	580	55.1	2500	12 ADN04443	Adn04443 Antipsori
16	580	55.1	2500	12 ADQ09273	Adq09273 Human RRM
17	580	55.1	2500	13 ACN37637	Acn37637 Tumour-as
18	580	55.1	3393	12 ADN03788	Adn03788 Antipsori
19	580	55.1	3393	12 ADO19225	Ado19225 Human PRO
20	578.4	54.9	1989	3 AAC78111	Aac78111 Human can

ALIGNMENTS

RESULT 1

AAF32447
ID AAF32447 standard; cDNA; 1053 BP.

XX
AC AAF32447;

DT 18-APR-2001 (first entry)

DE Human ribonucleotide reductase related nucleotide sequence SEQ ID NO:12.

XX Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.

OS Homo sapiens.

XX WO200100799-A1.

PD 04-JAN-2001.

PF 27-JUN-2000; 2000WO-JP004189.

XX 28-JUN-1999; 99JP-00181131.

PR 06-JUL-1999; 99JP-00192391.

PR 21-JAN-2000; 2000JP-00017770.

XX (TAKE) TAKEDA CHEM IND LTD.

PA (NAKA/) NAKAMURA Y.

PI Nakamura Y, Arakawa H, Tanaka H;

DR WPI; 2001-112446/12.

XX Ribonucleotide reductase involved in DNA repair and DNA encoding it, for

PT diagnosis, treatment and prevention of cancer.

XX Claim 6; Page 96-97; 102pp; Japanese.

PS The present invention describes a human ribonucleotide reductase
XX designated TP53R2H. TP53R2H has cytostatic activity. TP53R2H is part of
CC the DNA repair mechanism and its activity is induced by p53. It can be
CC used for the treatment, prevention and diagnosis of a wide range of
CC cancers. The present sequence represents a human ribonucleotide reductase
CC related sequence which is given in the exemplification of the present
CC invention

Adk70302 Respirato
Adj56536 Murine cD
Aah73325 Human cer
Aas44745 Human ful
Aav05641 Human rib
Ado57308 DNA enco
Aas41006 cDNA enco
Abi90228 Human pol
Aah07707 Human cDN
Aas79474 DNA enco
Abl14627 Drosophil
Aas62387 cDNA sequ
Aas51874 Human sec
Abt20895 Aspergill
Aas4564 Swinepox
Aas4948 Reverse c
Aaf84948 Nucleotid
Abl14626 Drosophil
AdS58378 Bacterial
Abt19075 Aspergill
Adt47814 Bacterial
AbS32250 Candida a
AdS49379 Bacterial
Acc61238 Gene sequ
Adk63241 Disease t

21 578.4 54.9 2215 12 ADK70302
22 578.4 54.9 2216 10 ADJ56536
23 578.4 54.9 3294 4 AAH73225
24 569 54.0 2641 4 AAS44745
25 568.8 54.0 1170 2 AAV05641
26 562.6 53.4 1328 12 ADO57308
27 547.8 52.0 978 4 AAS41006
28 545.4 51.8 977 6 ABL90228
29 527.2 50.1 608 4 AAH07707
30 514 48.8 1371 5 AAS79474
31 463 44.0 1289 4 ABL14627
32 442 42.0 1146 6 AAS62387
33 439.4 41.7 481 2 AAS51874
34 433.8 41.2 1218 8 ABT20895
35 429.8 40.8 14176 2 AAT84564
36 429.8 40.8 14176 4 AAF84948
37 429.8 40.8 14176 4 AAF84948
38 420.4 39.9 3945 4 ABL14626
39 404.8 38.4 963 13 ADS58378
40 398.6 37.9 1146 8 ABT19075
41 394.2 37.4 1173 13 ADT47814
42 393.6 37.4 1242 6 ABZ32250
43 388.2 36.9 1292 13 ADS49379
44 382 36.3 1200 10 ACC61238
45 382 36.3 1200 10 ADK63241

XX SQ Sequence 1053 BP; 314 A; 191 C; 246 G; 302 T; 0 U; 0 Other;

Query Match 100.0%; Score 1053; DB 4; Length 1053;
Best Local Similarity 100.0%; Pred. No. 9e-290;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGACCCCGAAGCGCGGCGGCTGGATCAGATGAGAGATCATCTTCA 60
DB 1 ATGGCGACCCCGAAGCGCGGCGGCTGGATCAGATGAGAGATCATCTTCA 60

QY 61 GACACCAACGAAAGTGAATAAAGTCAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
DB 61 GACACCAACGAAAGTGAATAAAGTCAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120

QY 121 CGTTTGTCTATCTTCCCAATCCAGTACCTGTATATTGGAATGTATATAACAGGCACAG 180
DB 121 CGTTTGTCTATCTTCCCAATCCAGTACCTGTATATTGGAATGTATATAACAGGCACAG 180

QY 181 GCTTCCTCTGACAGCAGAGAGGTGTGACTTATCAAGAGTCTCCCTCACTGGAAACAAG 240
DB 181 GCTTCCTCTGACAGCAGAGAGGTGTGACTTATCAAGAGTCTCCCTCACTGGAAACAAG 240

QY 241 CTTAAAGCAGATGAGAAGTACTTCTCTCATCTTACATCTTTAGCCCTTTTTCGACCCAGTAT 300
DB 241 CTTAAAGCAGATGAGAAGTACTTCTCTCATCTTACATCTTTAGCCCTTTTTCGACCCAGTAT 300

QY 301 GGAATGTAATGAATAATTTGTGAGCGCTTTAGTCAGGAGGTGAGTTCAGAGGCT 360
DB 301 GGAATGTAATGAATAATTTGTGAGCGCTTTAGTCAGGAGGTGAGTTCAGAGGCT 360

QY 361 CGCTGTTCTATGGCTTTCAAAATCTCATCGAGATGTTCACTCAGAGATGTACAGTTTG 420
DB 361 CGCTGTTCTATGGCTTTCAAAATCTCATCGAGATGTTCACTCAGAGATGTACAGTTTG 420

QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGTAA 480
DB 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGTAA 480

QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATTGGGCTTCGATGAGATAGCAGATAGAAA 540
DB 481 ACCATGCCCTATGTTAAGAAAAGCAGATTGGGCTTCGATGAGATAGCAGATAGAAA 540

QY 541 TCTACTTTTGGGAAGAGTGTGCGCTTGTGTGTAGAGAGTTCCTTCACGA 600
DB 541 TCTACTTTTGGGAAGAGTGTGCGCTTGTGTGTAGAGAGTTCCTTCACGA 600

QY 601 TCTTTTGTCTATATTTCTGGCTAAGAGAGAGAGTCTTATGCCAGGACTCACTTTTCC 660
DB 601 TCTTTTGTCTATATTTCTGGCTAAGAGAGAGAGTCTTATGCCAGGACTCACTTTTCC 660

QY 661 AATGAACATCAGCAGAGATGAAGACTTCTGTGACTTTGCTGCTGATGTTCCAA 720
DB 661 AATGAACATCAGCAGAGATGAAGACTTCTGTGACTTTGCTGCTGATGTTCCAA 720

QY 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGTGAGGAGATCATTTGTGATGCTGCA 780
DB 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGTGAGGAGATCATTTGTGATGCTGCA 780

QY 781 ATTGAGCAGGAGTTTAAACAGAAAGCCTTGGCAGTTGGCTCATTTGGAATGAATTCATT 840
DB 781 ATTGAGCAGGAGTTTAAACAGAAAGCCTTGGCAGTTGGCTCATTTGGAATGAATTCATT 840

QY 841 TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTAATTTGTGAACTTGGATTTCA 900
DB 841 TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTAATTTGTGAACTTGGATTTCA 900

QY 901 AAGGTTTTTCAGGCAGAAAATCCTTTGATTTTATGGAATAAATTTCTTTAGAGGAAA 960
DB 901 AAGGTTTTTCAGGCAGAAAATCCTTTGATTTTATGGAATAAATTTCTTTAGAGGAAA 960

QY 961 ACAAAATTTCTTGGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTTATGGCAGAAACC 1020

DB 961 ACAAAATTTCTTGGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020

QY 1021 ACAGATAAACGCTCTTCCCTTGGATGAGATTTT 1053
DB 1021 ACAGATAAACGCTCTTCCCTTGGATGAGATTTT 1053

RESULT 2
AAF32440
ID AAF32440 standard; cDNA; 1081 BP.
XX AAF32440;
AC AC
XX XX
DT 18-APR-2001 (first entry)
XX Human ribonucleotide reductase related nucleotide sequence SEQ ID NO:4.
DE Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.
KW Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.
XX Homo sapiens.
OS OS
PN WO200100799-A1.
XX 04-JAN-2001.
XX 27-JUN-2000; 2000WO-JP004189.
XX 28-JUN-1999; 99JP-00181131.
PR 06-JUL-1999; 99JP-00192391.
PR 21-JAN-2000; 2000JP-00017770.
XX (TAKE) TAKEDA CHEM IND LTD.
PA (NAKA/) NAKAMURA Y.
XX Nakamura Y, Arakawa H, Tanaka H;
PI WPI; 2001-112446/12.
XX Ribonucleotide reductase involved in DNA repair and DNA encoding it, for diagnosis, treatment and prevention of cancer.
PT Ribonucleotide reductase involved in DNA repair and DNA encoding it, for diagnosis, treatment and prevention of cancer.
XX Example 2; Page 91; 102pp; Japanese.
XX The present invention describes a human ribonucleotide reductase designated TP53R2H. TP53R2H has cytostatic activity. TP53R2H is part of the DNA repair mechanism and its activity is induced by p53. It can be used for the treatment, prevention and diagnosis of a wide range of cancers. The present sequence represents a human ribonucleotide reductase related sequence which is used in an example from the present invention
XX Sequence 1081 BP; 322 A; 199 C; 252 G; 308 T; 0 U; 0 Other;

Query Match 100.0%; Score 1053; DB 4; Length 1081;
Best Local Similarity 100.0%; Pred. No. 9.1e-290;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGACCCCGAAGCGCGGCGGCTGGATCAGATGAGAGATCATCTTCA 60
DB 20 ATGGCGACCCCGAAGCGCGGCGGCTGGATCAGATGAGAGATCATCTTCA 79

QY 61 GACACCAACGAAAGTGAATAAAGTCAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
DB 80 GACACCAACGAAAGTGAATAAAGTCAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 139

QY 121 CGTTTGTCTATCTTCCCAATCCAGTACCTGTATATTGGAATGTATATAACAGGCACAG 180
DB 140 CGTTTGTCTATCTTCCCAATCCAGTACCTGTATATTGGAATGTATATAACAGGCACAG 199

QY 181 GCTTCCTCTGAGCAGCAGAGAGGTGACTTATCAAGAGTCTCCCTCACTGGAAACAAG 240
DB 200 GCTTCCTCTGAGCAGCAGAGAGGTGACTTATCAAGAGTCTCCCTCACTGGAAACAAG 259

QY 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTTACATCTTTAGCCTTTTTCGACCCAGTAT 300

```
Db 260 CTTAAAGCAGATGAGAAGTACTTCACTCTCATCTAGCCTTTTTCAGCCAGTGAT 319
Qy 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTTCAGAGGCT 360
Db 320 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTTCAGAGGCT 379
Qy 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAGAAATGTTTCACTCAGAGATGTACAGTTG 420
Db 380 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAGAAATGTTTCACTCAGAGATGTACAGTTG 439
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATATGCAATGAA 480
Db 440 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATATGCAATGAA 499
Qy 481 ACCATGCCCTATGTTAAGAAAAGCAGATGGGCCCTTGGATGGATAGCAGATAGAAA 540
Db 500 ACCATGCCCTATGTTAAGAAAAGCAGATGGGCCCTTGGATGGATAGCAGATAGAAA 559
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGTCTGTGTAGAGGAGTCTTCTTCAGGA 600
Db 560 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGTCTGTGTAGAGGAGTCTTCTTCAGGA 619
Qy 601 TCTTTTGTCTATATTTCTGGCTTAAGAAAGAGAGTCTTATGCCAGACTCACTTTTCC 660
Db 620 TCTTTTGTCTATATTTCTGGCTTAAGAAAGAGAGTCTTATGCCAGACTCACTTTTCC 679
Qy 661 AATGAACATCATCAGCAGAGATGAGACTTCACTGTGACTTTTCTTGCCTGATGTTCCAA 720
Db 680 AATGAACATCATCAGCAGAGATGAGACTTCACTGTGACTTTTCTTGCCTGATGTTCCAA 739
Qy 721 TACTTAGTAAATAGCCTTCAGAGAAAGGTCAGGAGATCACTTGTATGCTGTCAAA 780
Db 740 TACTTAGTAAATAGCCTTCAGAGAAAGGTCAGGAGATCACTTGTATGCTGTCAAA 799
Qy 781 ATTGAGCAGGAGTCTTAAACAGAGCCTTGGCCAGTGGCTCATTTGGAATGAATTCATT 840
Db 800 ATTGAGCAGGAGTCTTAAACAGAGCCTTGGCCAGTGGCTCATTTGGAATGAATTCATT 859
Qy 841 TTGATGAAACAGTACATGAGTTTGTAGCTGACAGATTAATTTGGAACCTTGGATTTCA 900
Db 860 TTGATGAAACAGTACATGAGTTTGTAGCTGACAGATTAATTTGGAACCTTGGATTTCA 919
Qy 901 AAGGTTTTTTCAGCAGAAATCCTTTTGAATTTATGGAACAACTTTTTCAGAGGAAA 960
Db 920 AAGGTTTTTTCAGCAGAAATCCTTTTGAATTTATGGAACAACTTTTTCAGAGGAAA 979
Qy 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGATATCAGCGTTTTCAGTATGCGAGAAAC 1020
Db 980 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGATATCAGCGTTTTCAGTATGCGAGAAAC 1039
Qy 1021 ACAGATAAGCTCTTCACTTGGATGAGATTTT 1053
Db 1040 ACAGATAAGCTCTTCACTTGGATGAGATTTT 1072
```

RESULT 3

AAF32438
ID AAF32438 standard; cDNA; 1053 BP.

AC AAF32438;

XX AAF32438;
AC AAF32438;
XX AAF32438;
DT 18-APR-2001 (first entry)

DE Human ribonucleotide reductase TP53R2H nucleotide sequence SEQ ID NO:2.

XX Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.
OS Homo sapiens.

XX WO200100799-A1.

XX 04-JAN-2001.

```
XX 27-JUN-2000; 2000WO-JP004189.
XX
PR 28-JUN-1999; 99JP-00181131.
PR 06-JUL-1999; 99JP-00192391.
PR 21-JAN-2000; 2000JP-00017770.
XX
PA (TAKE ) TAKEDA CHEM IND LTD.
PA (NAKA/) NAKAMURA Y.
XX
PI Nakamura Y, Arakawa H, Tanaka H;
XX
DR WPI; 2001-112446/12.
DR P-PSDB; AAB69050.
XX
PT Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
PT diagnosis, treatment and prevention of cancer.
XX
PS Claim 5; Fig 1-3; 102pp; Japanese.
XX
CC The present sequence encodes a human ribonucleotide reductase designated
CC TP53R2H. TP53R2H has cytosolic activity. TP53R2H is part of the DNA
CC repair mechanism and its activity is induced by p53. It can be used for
CC the treatment, prevention and diagnosis of a wide range of cancers
XX
SQ Sequence 1053 BP; 314 A; 192 C; 246 G; 301 T; 0 U; 0 Other;
Query Match 99.8%; Score 1051.4; DB 4; Length 1053;
Best Local Similarity 99.9%; Pred. No. 2.6e-289;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGATCATCTTCA 60
Db 1 ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGATCATCTTCA 60
Qy 61 GACACCAACCAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAAGAAAGTCTCGC 120
Db 61 GACACCAACCAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAAGAAAGTCTCGC 120
Qy 121 CGGTTTGTCTATCTTCCAAATCCAGTACCTGATATTTGGAAAATGTATAAACAGGCACAG 180
Db 121 CGGTTTGTCTATCTTCCAAATCCAGTACCTGATATTTGGAAAATGTATAAACAGGCACAG 180
Qy 181 GCTTCTCTTGGACAGAGAGGTTGACTTATCAAGAGATCTCCCTCAGTGGAAACAAG 240
Db 181 GCTTCTCTTGGACAGAGAGGTTGACTTATCAAGAGATCTCCCTCAGTGGAAACAAG 240
Qy 241 CTTAAAGCAGATGAGAAGTACTTCTCATCATCTTTAGCCTTTTTCAGCCAGTGTAT 300
Db 241 CTTAAAGCAGATGAGAAGTACTTCTCATCATCTTTAGCCTTTTTCAGCCAGTGTAT 300
Qy 301 GGAATTTGTAATGAAATTTTGGTGGAGCGCTTTAGTCAGGAGGTGAGGCTTCAGAGGCT 360
Db 301 GGAATTTGTAATGAAATTTTGGTGGAGCGCTTTAGTCAGGAGGTGAGGCTTCAGAGGCT 360
Qy 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCAGAAATGTTTCACTCAGAGATGTACAGTTG 420
Db 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCAGAAATGTTTCACTCAGAGATGTACAGTTG 420
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATATGCAATGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATATGCAATGAA 480
Qy 481 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGGCCCTTGGATGGATAGCAGATAGAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGGCCCTTGGATGGATAGCAGATAGAAA 540
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTCTGTAGAGGAGTCTTCTTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTCTGTAGAGGAGTCTTCTTCAGGA 600
Qy 601 TCTTTTGTCTATATTTCTGGCTTAAGAAAGAGAGTCTTATGCCAGACTCACTTTTCC 660
```


Db 601 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGCTTTATGCCAGGACTCACTTTTTC 660
 QY 661 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
 Db 661 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
 QY 721 TACTTTAGTAATAAGCTTTTCAAGAAAGGGTCAAGGAGATCATTTTGTGATGCTGTCAA 780
 Db 721 TACTTTAGTAATAAGCTTTTCAAGAAAGGGTCAAGGAGATCATTTTGTGATGCTGTCAA 780
 QY 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGGTTGGCTCATTTGGAATGAATTGCAAT 840
 Db 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGGTTGGCTCATTTGGAATGAATTGCAAT 840
 QY 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTACTTTGTGAACTTGGATTTCTCA 900
 Db 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTACTTTGTGAACTTGGATTTCTCA 900
 QY 901 AAGGTTTTTTCAGGCAGAAAATCCTTTTGTATTTTATGGAACAACTTTCTTTAGAGGAAAA 960
 Db 901 AAGGTTTTTTCAGGCAGAAAATCCTTTTGTATTTTATGGAACAACTTTCTTTAGAGGAAAA 960
 QY 961 ACAAAATTTCTTGAGAAACAGGTTTCAGAGTATCAGGTTTTCAGATTATGCGAGAAACC 1020
 Db 961 ACAAAATTTCTTGAGAAACAGGTTTCAGAGTATCAGGTTTTCAGATTATGCGAGAAACC 1020
 QY 1021 ACAGATAAGCTTTTACCTTGGATGTCAGATT 1053
 Db 1021 ACAGATAAGCTTTTACCTTGGATGTCAGATT 1053

RESULT 4

AAAL2411

ID AAAL2411 standard; cDNA; 2596 BP.

XX AC AAAL2411;

XX DT 25-JUL-2000 (first entry)

XX DE cDNA encoding a human RNA-associated protein.

XX KW Human; RNA-associated protein; cell proliferation; cancer; inflammation;
 KW immune response; reproductive disorder; actinic keratosis;
 KW atherosclerosis; arteriosclerosis; bursitis; cirrhosis; hepatitis;
 KW mixed connective tissue disease; myelofibrosis; primary thrombocythemia;
 KW paroxysmal nocturnal hemoglobinuria; polycythemia vera; psoriasis;
 KW trauma; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT CDS 56..1111

XX FT /*tag= a

XX FT /product= "RNA-associated protein"

XX PN WO200015799-A2.

XX PD 23-MAR-2000.

XX PF 17-SEP-1999; 99WO-US021688.

XX PR 17-SEP-1998; 98US-00156039.

XX PR 22-SEP-1998; 98US-00158720.

XX PR 04-NOV-1998; 98US-00186815.

XX PR 08-APR-1999; 99US-0128660P.

XX PA (INCY-) INCYTE PHARM INC.

XX PI Tang YT, Corley NC, Guegler KJ, Gorgone GA, Patterson C;

XX PI Hillman JL, Baughn MR, Lal P, Azimzai Y, Yue H, Yang J;

XX DR WPI; 2000-271437/23.

XX DR P-PSDB; AAY84439.

XX New polypeptides and polynucleotides, useful for preventing and treating
 PT a disorder associated with increased or decreased expression of RNA
 PT associated proteins.

XX PS Claim 9; Page 120-121; 131pp; English.

XX The present sequence encodes a human RNA-associated protein. The
 CC expression of RNA-associated proteins is closely associated with
 CC reproductive tissues, nervous tissues, cell proliferation including
 CC cancer, inflammation and immune responses, and so they may be used for
 CC diagnosis, treatment or prevention of cell proliferative,
 CC immune/inflammatory disorders, and reproductive disorders. Diseases and
 CC disorders which may be treated include actinic keratosis,
 CC atherosclerosis, arteriosclerosis, bursitis, cirrhosis, hepatitis, mixed
 CC connective tissue disease, myelofibrosis, paroxysmal nocturnal
 CC hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia
 CC and cancers, and trauma

XX SQ Sequence 2596 BP; 783 A; 418 C; 510 G; 885 T; 0 U; 0 Other;

Query Match 99.8%; Score 1051.4; DB 3; Length 2596;

Best Local Similarity 99.9%; Pred. No. 3.9e-289;

Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGCACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60

Db 56 ATGGCGCACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 115

QY 61 GACACCAACGAAGTGAATAAAGTCAAAATGAAGAGCCACTCCTAAGAAAGAGTTCTCG 120

Db 116 GACACCAACGAAGTGAATAAAGTCAAAATGAAGAGCCACTCCTAAGAAAGAGTTCTCG 175

QY 121 CGGTTTGTCTATCTTTTCCAAATCCAGTACCTGATATTTGGAAATATATAAAGGACAG 180

Db 176 CGGTTTGTCTATCTTTTCCAAATCCAGTACCTGATATTTGGAAATATATAAAGGACAG 235

QY 181 GCTTCTCTTGGACAGCAGAGAGGTTGACTTATCAAAAGGATCTCCCTCAGTGGAAACAG 240

Db 236 GCTTCTCTTGGACAGCAGAGAGGTTGACTTATCAAAAGGATCTCCCTCAGTGGAAACAG 295

QY 241 CTTAAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTGTAT 300

Db 296 CTTAAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTGTAT 355

QY 301 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTTCAGGAGGTCAGGTTCCAGAGGCT 360

Db 356 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTTCAGGAGGTCAGGTTCCAGAGGCT 415

QY 361 CGCTGTTTCTATGCTTCAAAATCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG 420

Db 416 CGCTGTTTCTATGCTTCAAAATCTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG 475

QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAATCAATTGAA 480

Db 476 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAATCAATTGAA 535

QY 481 ACCATGCCCTTATGTTAAGAAAAAGCAGATTGGGCTTTGCGATGGATGACAGATGAAAA 540

Db 536 ACCATGCCCTTATGTTAAGAAAAAGCAGATTGGGCTTTGCGATGGATGACAGATGAAAA 595

QY 541 TCTACTTTGGGGAAGAGTGGTGGCCCTTTCTGCTGTAGAGAGTTCCTTCTCAGGA 600

Db 596 TCTACTTTGGGGAAGAGTGGTGGCCCTTTCTGCTGTAGAGAGTTCCTTCTCAGGA 655

QY 601 TCTTTTGTCTGTATTTCTGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTCC 660

Db 656 TCTTTTGTCTGTATTTCTGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTCC 715

QY 661 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720

Db 716 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 775

QY 721 TACTTAGTAAATAAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGTTGATGCTGTCAAA 780
DB 776 TACTTAGTAAATAAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGTTGATGCTGTCAAA 835
QY 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTTGGCTCATTTGGAATGAATTGCATT 840
DB 836 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTTGGCTCATTTGGAATGAATTGCATT 895
QY 841 TTGATGAAACAGTACATTGAGTTTGTAGCTGTGACAGATTACTTTGTGAACTTTGATTTCTCA 900
DB 896 TTGATGAAACAGTACATTGAGTTTGTAGCTGTGACAGATTACTTTGTGAACTTTGATTTCTCA 955
QY 901 AAGGTTTTTTCAGGAGAAATCCTTTTGAATTTATGAAAAACATTTCTTTAGAGGAAAA 960
DB 956 AAGGTTTTTTCAGGAGAAATCCTTTTGAATTTATGAAAAACATTTCTTTAGAGGAAAA 1015
QY 961 ACAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGATTATGCGAGAAACC 1020
DB 1016 ACAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGATTATGCGAGAAACC 1075
QY 1021 ACAGATAAGCTCTTACCTTTGGATGCAGATTTT 1053
DB 1076 ACAGATAAGCTCTTACCTTTGGATGCAGATTTT 1108

RESULT 5
AAF32439
ID AAF32439 standard; cDNA; 4955 BP.
XX AC AAF32439;
XX DT 18-APR-2001 (first entry)
DE Human ribonucleotide reductase related nucleotide sequence SEQ ID NO:3.
XX Human; ribonucleotide reductase; cancer; DNA repair; p53; ss.
XX OS Homo sapiens.
XX PN WO200100799-A1.
XX PD 04-JAN-2001.
XX PF 27-JUN-2000; 2000WO-JP004189.
XX PR 28-JUN-1999; 95JP-00181131.
XX PR 06-JUL-1999; 95JP-00192391.
XX PR 21-JAN-2000; 2000JP-00017770.
XX PA (TAKE) TAKEDA CHEM IND LTD.
XX PI (NAKA/) NAKAMURA Y.
XX PI Nakamura Y, Arakawa H, Tanaka H;
XX WPI; 2001-112446/12.
XX DR
XX PT Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
XX PT diagnosis, treatment and prevention of cancer.
XX PS
XX PS Example 2; Page 87-90; 102pp; Japanese.
CC The present invention describes a human ribonucleotide reductase
CC designated TP53R2H. TP53R2H has cytosolic activity. TP53R2H is part of
CC the DNA repair mechanism and its activity is induced by p53. It can be
CC used for the treatment, prevention and diagnosis of a wide range of
CC cancers. The present sequence represents a human ribonucleotide reductase
CC related sequence which is used in an example from the present invention
SQ Sequence 4955 BP; 1526 A; 803 C; 999 G; 1627 T; 0 U; 0 Other;
Query Match 99.8%; Score 1051.4; DB 4; Length 4955;
Best Local Similarity 99.9%; Pred. No. 5.2e-289;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGACCCGGAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
DB 245 ATGGCGACCCGGAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
QY 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
DB 305 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 364
QY 121 CGGTTTGTCTATCTTCCAAATCCAGTACCCCTGATATTTGGAAATGTATAACAGGCACAG 180
DB 365 CGGTTTGTCTATCTTCCAAATCCAGTACCCCTGATATTTGGAAATGTATAACAGGCACAG 424
QY 181 GCTTCTCTTGGACAGCAGAGAGTTGACTTATCAAAAGAGTCTCCCTCACTTGGAAACAG 240
DB 425 GCTTCTCTTGGACAGCAGAGAGTTGACTTATCAAAAGAGTCTCCCTCACTTGGAAACAG 484
QY 241 CTTAAAGCAGATGAGAGTACTTTCATCTCTCACATCTTAGCCCTTTTTCAGCCAGATGAT 300
DB 485 CTTAAAGCAGATGAGAGTACTTTCATCTCTCACATCTTAGCCCTTTTTCAGCCAGATGAT 544
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGAGTGCAGGTTCCAGAGGCT 360
DB 545 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGAGTGCAGGTTCCAGAGGCT 604
QY 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAATGTTTCACTCAGAGATGTACAGTTTG 420
DB 605 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAATGTTTCACTCAGAGATGTACAGTTTG 664
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAAATGCAATTGAA 480
DB 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAAATGCAATTGAA 724
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCTTTCGATGATGATAGAGATAAGAAA 540
DB 725 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCTTTCGATGATGATAGAGATAAGAAA 784
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGAGTTTCTTCTCAGGA 600
DB 785 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGAGTTTCTTCTCAGGA 844
QY 601 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGTCTTATGCCAGAGACTCACATTTTTC 660
DB 845 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGTCTTATGCCAGAGACTCACATTTTTC 904
QY 661 AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTGTGCTGATGTTCCAA 720
DB 905 AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTGTGCTGATGTTCCAA 964
QY 721 TACTTAGTAAATAAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGTTGATGCTGTCAAA 780
DB 965 TACTTAGTAAATAAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGTTGATGCTGTCAAA 1024
QY 781 ATTGAGCAGGAGTTTAAACAGAGCCTTCCAGTTGGCCTCATTTGGAATGAATTGCATT 840
DB 1025 ATTGAGCAGGAGTTTAAACAGAGCCTTCCAGTTGGCCTCATTTGGAATGAATTGCATT 1084
QY 841 TTGATGAAACAGTACATTGAGTTTGTAGCTGTGACAGATTACTTTGGAATTTGATTTCTCA 900
DB 1085 TTGATGAAACAGTACATTGAGTTTGTAGCTGTGACAGATTACTTTGGAATTTGATTTCTCA 1144
QY 901 AAGGTTTTTTCAGGAGAAATCCTTTTGAATTTTATGAAAAACATTTCTTTAGAGGAAAA 960
DB 1145 AAGGTTTTTTCAGGAGAAATCCTTTTGAATTTTATGAAAAACATTTCTTTAGAGGAAAA 1204
QY 961 ACAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGATTATGCGAGAAACC 1020
DB 1205 ACAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGATTATGCGAGAAACC 1264
QY 1021 ACAGATAAGCTCTTACCTTTGGATGCAGATTTT 1053
DB 1265 ACAGATAAGCTCTTACCTTTGGATGCAGATTTT 1297

RESULT 6
 ID ADR24210 standard; DNA; 4955 BP.
 AC ADR24210;
 DT 21-OCT-2004 (first entry)
 XX Breast cancer prognosis marker #71.
 XX ds; breast cancer; prognosis; gene expression; diagnosis.
 KW Homo sapiens.
 OS WO2004065545-A2.
 PN 05-AUG-2004.
 XX 15-JAN-2004; 2004WO-US001100.
 XX 15-JAN-2003; 2003US-00342887.
 PA (ROSE-) ROSETTA INPHARMATICS LLC.
 PA (NECA-) NETHERLANDS CANCER INST.
 XX Van't Veer LJ, He Y;
 XX WPI; 2004-593473/57.
 DR
 CC Classifying a breast cancer patient according to prognosis comprises
 CC determining the similarity between the level of expression of each of
 CC five genes in a cell sample taken from patient, to control levels.
 PT
 PT Disclosure; SEQ ID NO 71; 226pp; English.
 PS
 CC The invention relates to a method of classifying a breast cancer patient
 CC according to prognosis by determining the similarity between the level of
 CC expression of each of five genes for which markers are listed in the
 CC specification, in a cell sample taken from the breast cancer patient, to
 CC control levels of expression for each respective five genes to obtain a
 CC patient similarity value. The methods are useful for classifying a breast
 CC cancer patient according to prognosis. Kits and computer program products
 CC are useful for data analysis using the diagnostic, prognostic and
 CC statistical methods of the invention. This sequence corresponds to a
 CC marker used in the method of the invention.
 XX
 SQ Sequence 4955 BP; 1526 A; 803 C; 999 G; 1627 T; 0 U; 0 Other;
 Query Match 99.8%; Score 1051.4; DB 13; Length 4955;
 Best Local Similarity 99.9%; Pred. No. 5.2e-289;
 Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGATCATCTTCA 60
 DB 245 ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGATCATCTTCA 304
 QY 61 GACACCAACGAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAGAGTTCTCGC 120
 DB 305 GACACCAACGAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAGAGTTCTCGC 364
 QY 121 CGTTTGTTCATCTTCCCAATCCAGTACCTGATATTTGGAATGATATAACAGGCACAG 180
 DB 365 CGTTTGTTCATCTTCCCAATCCAGTACCTGATATTTGGAATGATATAACAGGCACAG 424
 QY 181 GCTTCTCTTGACAGCAGAGAGGTTGACTTATCAAGGATCTCCCTCACTGGAACAAG 240
 DB 425 GCTTCTCTTGACAGCAGAGAGGTTGACTTATCAAGGATCTCCCTCACTGGAACAAG 484
 QY 241 CTTAAAGCAGATGAGAAGTACTTCTCTCTCATCTTACATCTTTTTCAGCCAGTGAT 300
 DB 485 CTTAAAGCAGATGAGAAGTACTTCTCTCTCATCTTACATCTTTTTCAGCCAGTGAT 544

QY 301 GGAATCTTAATGAAATTTTGGTGGAGCGCTTTAGTCAGGAGGTGAGGTTCCAGAGGCT 360
 DB 545 GGAATCTTAATGAAATTTTGGTGGAGCGCTTTAGTCAGGAGGTGAGGTTCCAGAGGCT 604
 QY 361 CGCTGTTTCTATGCGCTTTCAAATTTCTCATCGAATGTTCTACTCAGAGATGTACAGTTTG 420
 DB 605 CGCTGTTTCTATGCGCTTTCAAATTTCTCATCGAATGTTCTACTCAGAGATGTACAGTTTG 664
 QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAAATCAATTTGAA 480
 DB 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAAATCAATTTGAA 724
 QY 481 ACCATGCCCTATGTTAAAGAAAAAGCAGATTGGGCTTTGCGATGATAGCAGATAGAAA 540
 DB 725 ACCATGCCCTATGTTAAAGAAAAAGCAGATTGGGCTTTGCGATGATAGCAGATAGAAA 784
 QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAAAGGAGTTTCTTCTCAGGA 600
 DB 785 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAAAGGAGTTTCTTCTCAGGA 844
 QY 601 TCTTTTGTCTATATTCTGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTCTCC 660
 DB 845 TCTTTTGTCTATATTCTGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTCTCC 904
 QY 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGGCTGATGTTCCAA 720
 DB 905 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGGCTGATGTTCCAA 964
 QY 721 TACTTAGTAATAAGCCTTTCAGAAAGAGGTCAAGGAGATCATTTGTTGATGCTGTCAAA 780
 DB 965 TACTTAGTAATAAGCCTTTCAGAAAGAGGTCAAGGAGATCATTTGTTGATGCTGTCAAA 1024
 QY 781 ATTGAGCAGGAGTTTAAACAGAGAGCCTTGCAGTGGCCTCATTTGGAATGAATTCATT 840
 DB 1025 ATTGAGCAGGAGTTTAAACAGAGAGCCTTGCAGTGGCCTCATTTGGAATGAATTCATT 1084
 QY 841 TTGATGAAACAGTACATTTAGCTTTGTAGCTCAGACAGATTACTTGTGAACTTCGATCTCA 900
 DB 1085 TTGATGAAACAGTACATTTAGCTTTGTAGCTCAGACAGATTACTTGTGAACTTCGATCTCA 1144
 QY 901 AAGGTTTTTCAAGCAGAAAAATCCTTTTGAATTTTATGGAACAACTTTCTTTAGAGAAAA 960
 DB 1145 AAGGTTTTTCAAGCAGAAAAATCCTTTTGAATTTTATGGAACAACTTTCTTTAGAGAAAA 1204
 QY 961 ACAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC 1020
 DB 1205 ACAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC 1264
 QY 1021 ACAGATAACGTCCTTCACTTGGATGACAGATTTT 1053
 DB 1265 ACAGATAACGTCCTTCACTTGGATGACAGATTTT 1297
 RESULT 7
 AAHL4924
 ID AAHL4924 standard; cDNA; 1601 BP.
 XX
 AC AAHL4924;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human cDNA sequence SEQ ID NO:12810.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 OS Homo sapiens.
 XX EF1074617-A2.
 PN
 XX 07-FEB-2001.
 PD
 XX 28-JUL-2000; 2000EP-00116126.
 PF
 XX

29-JUL-1999; 99JP-00248036.
 27-AUG-1999; 99JP-00300253.
 11-JAN-2000; 2000JP-00118776.
 02-MAY-2000; 2000JP-00183767.
 09-JUN-2000; 2000JP-00241899.
 (HELI-) HELIX RES INST.
 Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 WPI; 2001-318749/34.
 Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
 Claim 8; SEQ ID NO 12810; 2537pp + Sequence Listing; English.
 The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH0166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent polynucleotides, all of which are used in the exemplification of the present invention.
 Query Match 99.5%; Score 1048.2; DB 4; Length 1601;
 Best Local Similarity 99.7%; Pred. No. 2.5e-288;
 Matches 1050; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 1 ATGGCGGACCCGGAAGCCGGAAGCGCGCGCTGGATCAGATGAGATCACTTCA 60
 40 ATGGCGGACCCGGAAGCCGGAAGCGCGCGCTGGATCAGATGAGATCACTTCA 99
 61 GACACCAACGAAAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 120
 100 GACACCAACGAAAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 159
 121 CGTTTGTTCATCTTCCATCCAGTACCCTGATATTGGAAATGTATAAAGGACAG 180
 160 CGTTTGTTCATCTTCCATCCAGTACCCTGATATTGGAAATGTATAAAGGACAG 219
 181 GCTTCCTTCTGACAGCAGAGAGGTTGATTCATCAAGGATCTCCCTCACTGGACAA 240
 220 GCTTCCTTCTGACAGCAGAGAGGTTGATTCATCAAGGATCTCCCTCACTGGACAA 279
 241 CTTAAGCAGATGAGAAGTACTTCTCTCATCTCTAGCCCTTTTTCAGCCAGTGAT 300
 280 CTTAAGCAGATGAGAAGTACTTCTCTCATCTCTAGCCCTTTTTCAGCCAGTGAT 339
 301 GGAATTTGTAATGAAAATTTTGGTGGAGCGCTTTAGTCAGGAGGTCCAGAGGCT 360
 340 GGAATTTGTAATGAAAATTTTGGTGGAGCGCTTTAGTCAGGAGGTCCAGAGGCT 399

QY 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCAGAGATGTTCTACTCAGAGATGTACAGTTTG 420
 DB 400 CGCTGCTTCTATGGCTTTCAAAATTTCTCATCAGAGATGTTCTACTCAGAGATGTACAGTTTG 459
 QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATCAATGAA 480
 DB 460 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATCAATGAA 519
 QY 481 ACCATGCCCTATGTTAAGAAAAGGAGATGGGCTTTGCCGATGATGATGATGATGATGATGAT 540
 DB 520 ACCATGCCCTATGTTAAGAAAAGGAGATGGGCTTTGCCGATGATGATGATGATGATGATGAT 579
 QY 541 TCTACTTTTGGGAAAAGAGTGGTGGCTTTCTGCTGTAGAGAGGAGTTTCTTCTCTCAGCA 600
 DB 580 TCTACTTTTGGGAAAAGAGTGGTGGCTTTCTGCTGTAGAGAGGAGTTTCTTCTCTCAGCA 639
 QY 601 TCTTTTGTCTCTATATTCTGGCTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTCTCC 660
 DB 640 TCTTTTGTCTCTATATTCTGGCTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTCTCC 699
 QY 661 AATGAACTCATCAGCAGAGATGAGGACTTCACTGTGACTTTGCTGCTGATGTTCCAA 720
 DB 700 AATGAACTCATCAGCAGAGATGAGGACTTCACTGTGACTTTGCTGCTGATGTTCCAA 759
 QY 721 TACTTAGTAATAAGCCCTTCAGAAAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 780
 DB 760 TACTTAGTAATAAGCCCTTCAGAAAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 819
 QY 781 ATTGAGCAGGAGTTTAAACAGAAAGCCCTTGCAGTGGCCCTCATTTGGAATGAATGCAATT 840
 DB 820 ATTGAGCAGGAGTTTAAACAGAAAGCCCTTGCAGTGGCCCTCATTTGGAATGAATGCAATT 879
 QY 841 TTGATGAACAGTACATAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 900
 DB 880 TTGATGAACAGTACATAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 939
 QY 901 AAGSTTTTTCAGGCGAGAAAATCCCTTTGATTTTATGAAAACATTTCTTTAGAGGAAA 960
 DB 940 AAGSTTTTTCAGGCGAGAAAATCCCTTTGATTTTATGAAAACATTTCTTTAGAGGAAA 999
 QY 961 ACAAAATTTCTTTCAGAAACCGAGTTTCAGAGTATCAGCGTTTTCAGAGTTATGCGAGAAACC 1020
 DB 1000 ACAAAATTTCTTTCAGAAACCGAGTTTCAGAGTATCAGCGTTTTCAGAGTTATGCGAGAAACC 1059
 QY 1021 ACAGATAACCTCTTCACTTGGATGCGAGATTTT 1053
 DB 1060 ACAGATAACCTCTTCACTTGGATGCGAGATTTT 1092
 RESULT 8
 ADQ57092
 ID ADQ57092 standard; DNA; 706 BP.
 AC ADQ57092;
 XX
 AC ADQ57092;
 XX
 DT 21-OCT-2004 (first entry)
 XX
 DE Novel canine microarray-related DNA sequence SeqID8394.
 XX
 DE canine microarray; drug screening; toxicity assay;
 KW environmental pollutant; cellular response; gene expression profile;
 KW toxic response; liver necrosis; fatty liver disease;
 KW protein adduct formation; hepatitis; dog; ds.
 XX
 OS Canis familiaris.
 XX
 XX WO2004063324-A2.
 PN
 XX
 PD 29-JUL-2004.
 XX
 XX 05-MAY-2003; 2003WO-US013853.
 XX
 XX 03-MAY-2002; 2002US-0377240P.
 PR

(GENE-) GENE LOGIC INC.
(PFIZ) PFIZER PROD INC.

Diggins JC, Porter M, Wei T;
WPI; 2004-561890/54.

New isolated nucleic acid molecule, useful for drug screening and toxicity assays or for assessing the impact, including toxicity, of a compound, pharmaceutical agent or environmental pollutant on a cell or living organism.

Claim 1; SEQ ID NO 8394; 41pp; English.

This invention is related to a novel isolated canine nucleic acid sequences and the construction of canine microarrays containing a significant portion of the canine genome. The isolated canine nucleic acid sequences of the invention may be useful for drug screening and toxicity assays. The invention is therefore useful for assessing the impact, including toxicity, of a compound, pharmaceutical agent or environmental pollutant on a cell or living organism. The methods are useful for detecting genes that are up- or down-regulated in canines in a disease state. The sequences are useful as diagnostic agents or markers to detect a cellular response in a sample individually or as part of a gene expression profile. It is also useful as a target for agents that modulate gene expression or activity. The database is useful for producing electronic Northernblots that allow the user to determine the cell type or tissue in which a given gene is expressed and to allow determination of the abundance or expression level of a given gene in a particular tissue or cell. The methods are useful for determining the similarity of a toxic response to one or more individual compounds. The methods are useful for predicting at least one toxic response or the likelihood that a compound or test agent will induce various specific pathologies such as those of the liver (liver necrosis, fatty liver disease, protein adduct formation or hepatitis), those of the kidney, heart, brain or testes, or other pathologies associated with at least one of the toxins. The methods are also useful for predicting or elucidating the potential cellular pathways influenced, induced or modulated by the compound or test agent due to the similarity of the expression profile compared to the profile induced by a known toxin. The present sequence is that of a canine DNA sequence which was claimed for use during the production of a canine microarray of the invention.

Sequence 706 BP; 208 A; 113 C; 158 G; 227 T; 0 U; 0 Other;

Query Match 61.8%; Score 650.6; DB 13; Length 706;
Best Local Similarity 95.2%; Pred. No. 5.5e-175;
Matches 671; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

349 GTTCCAGAGGCTCGGTGTTCTATCGGCTTCAAATTTCTCATCGAGATGTTCACTCAGAG 408
1 GTTCCAGAGGCTCGGTGTTCTATCGGCTTCAAATTTCTCATCGAGATGTTCACTCAGAG 60

409 ATGTACAGTTTCTGATAGACACTTACATCAGAGATCCCAAGAAAGGAATTTTATTT 468
61 ATGTACAGTTTCTGATAGACACTTACATCAGAGATCCCAAGAAAGGAATTTTATTT 120

469 AATGCAATTTGAACCATGCTTCTTAAGAAAGAGAGATGGGCTTCGATGATA 528
121 AATGCAATTTGAACCATGCTTCTTAAGAAAGAGAGATGGGCTTCGATGATA 180

529 CGAGATAGAAAATCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAAGGAGTT 588
181 GAAGATAGAAAATCTACTTTTGGGAAAGAGATGCTGGCTTTTGTCTGTAGAAGGAGTT 240

589 TTCTTCTCAGGATCTTTTGTCTCTATATTTGGCTTAAAGAGAGAGGCTTTATGCCAGGA 648
241 TTCTTCTCAGGATCTTTTGTCTCTATATTTGGCTTAAAGAGAGAGGCTTCTATGCCCTGA 300

649 CTCACATTTTCAATGAATCTATCAGCAGAGATGAAGCACTTCACTGTGACTTTGCTTGC 708
301 CTCACATTTTCAATGAATCTATCAGCAGAGATGAAGGCTTCACTGTGACTTTGCTTGC 360

709 CTGATGTTTCCAATACCTTAAATAAGCCTTCAGAAAGAAAGGGTCAGGAGATCAATGTT 766
361 CTGATGTTTCCAATACCTTAAATAAGCCTTCAGAAAGAAAGGGTCAGGAGATCAATGTT 420

769 GATGCTGCTCAAAATTCAGCAGGAGTTTAAACAGAGAGCCTTCGCCAGTTGGCCCTCATTTGA 828
421 AATGCTGTTTGAATAATGAGCAGGAGTTTAAACAGAGAGCCTTCGCCAGTTGGCCCTCATTTGA 480

829 ATGAATGTCATTTTGTATGAAACAGTACATTTAGTTTGTAGCTGCACAGATTAATCTTGTGGA 888
481 ATGAATGTCATTTTGTATGAAACAGTATATTTAGTTTGTAGCTGCACAGATTAATCTTGTGGA 540

889 CTTGGATCTCAAAAGGTTTTCAGGCAGAAAATCCCTTTGATTATTTATGAAACCAATTTCT 948
541 CTTGGATCTCAAAAGGTTTTCAGGCAGAAAATCCCTTTGATTATTTATGAAACCAATTTCT 600

949 TTAGAGGAGAAACAAATTTCTTTCAGAAACAGTTTTCAGAGTATCAGCGTTTTCGAGTT 1000
601 TTAGAGGAGAAACAAATTTCTTTCAGAAACAGTTTTCAGAGTATCAGCGTTTTCGCGTT 660

1009 ATGCGAGAAACACAGATAACGTTCTTCACTTGGATGCAGATTTT 1053
661 ATGCGAGAAACACAGATAATGTTCTTCACTTGGATGCAGATTTT 705

RESULT 9
ID AAS44917/c
AA AAS44917 standard; DNA; 2482 BP.
XX AA AAS44917;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human contig polynucleotide sequence #170.
XX
XX Mammal; human; rhesus monkey; baker's yeast; fission yeast; Norway rat;
XX mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;
XX cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;
XX nervous system disorder; inflammatory disorder; cell differentiation; ds;
XX angiogenesis; stem cell growth factor; activin; inhibin; cartilage; burn;
XX genetic disorder; bone regeneration; tendon; ligament; tissue repair;
XX cytoskeletal; antirheumatic; antiarthritic; vulnary; antiinflammatory;
XX antibacterial; immunosuppressive; vasotropic; antiparkinsonian;
XX neuroprotective; osteopathic; antidiabetic; antiasthmatic; antiallergic;
XX immunostimulant; analgesic; gene therapy.

XX Homo sapiens.
OS Synthetic.
XX
XX WO200164834-A2.
XX
XX 07-SEP-2001.
XX
XX 26-FEB-2001; 2001WO-US004926.
XX
XX 28-FEB-2000; 2000US-00515126.
XX
XX 18-MAY-2000; 2000US-00577409.
XX
XX 17-JUN-2000; 2000US-00597707.
XX
XX 14-JUL-2000; 2000US-00616807.
XX
XX 19-SEP-2000; 2000US-00664641.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
XX Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;
XX Drmanac R;
XX
XX WPI; 2001-589862/66.
XX
XX P-PSDB; AAU28017.
XX
XX Novel polypeptides and nucleic acids obtained from cDNA libraries
XX prepared from various human tissues, for diagnosis, treatment of cancer,
XX

PT neurological, inflammatory disorders and for use in arrays for detection.

PS Claim 1; SEQ ID NO 514; 153pp; English.

XX
XX
CC Sequences AAS44576-AAS44919 represent full-length polynucleotides and
CC contig polynucleotides encoding polypeptides of the invention. The DNA
CC and protein sequences are useful for the treatment, diagnosis and
CC prevention of various types of disorder in a mammalian subject such as a
CC human, dog, monkey, mouse, hamster or rat. The disorders include cancers
CC such as leukaemia, lymphoma and neuroblastoma, autoimmune disorders such
CC as multiple sclerosis, connective tissue disease, rheumatoid arthritis,
CC diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system
CC disorders such as Parkinson's disease, Alzheimer's disease, Huntington's
CC chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and
CC Wernicke disease, inflammatory disorders such as nephritis, Crohn's
CC disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory
CC bowel disease. The sequences exhibit activity relating to angiogenesis,
CC cell proliferation, cell differentiation, stem cell growth factor,
CC activin or inhibin. Therefore, they can be used to manipulate stem cells
CC in culture to give rise to neuroepithelial cells that can be used to
CC augment or replace cells damaged by illness, accidental damage or genetic
CC disorders. The sequences may also be used for regeneration of bone,
CC cartilage, tendons and ligaments and in tissue repair and burn healing.
CC Note: Some sequences for this patent did not form part of the printed
CC specification, but were obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 2482 BP; 707 A; 598 C; 522 G; 655 T; 0 U; 0 Other;

Query Match 55.1%; Score 580; DB 4; Length 2482;

Best Local Similarity 75.1%; Pred. No. 1.4e-154;

Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

QY	89	ATGAAGAGCCATCTCTAAGAAAGATTCTCGCGGTTTGTCATCTTCCCAATCCAGTACC	148
DB	2086	AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTGTCTATCTTCCCATCGAGTACC	2027
QY	149	CTGATATTGGAAATGTAACAGGACAGGCTTCCTTGGACAGCAGAGAGGTG	208
DB	2026	ATGATATCTGGCAGATGTAATAGAGGCGCAGAGGCTTCTTTGGACCGCGAGAGGTG	1967
QY	209	ACTTATCAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAAGTACTTCATCT	268
DB	1966	ACCTCTCCAGGACATTCAGACTGGGAATCCCTGAAACCCGGAGAGATATTTATAT	1907
QY	269	CTCACATCTTAGCTTTTTCAGCAGCAGTGATGGAATTTGATGAAATTTGGTGAGC	328
DB	1906	CCCATGTTCTGGCTTCTTTTGCAGCAAGCATGGCATAGTAATGAACACTTGGTGAGC	1847
QY	329	GCTTTAGTCAGAGGTGAGGTTCCAGAGGTCGCTGTTTCTATGCGCTTCAAAATCTCA	388
DB	1846	GATTTAGCCAAAGTTCAGATTACAGAGCCGCTGTTTCTATGCTTCCAAATTGCCA	1787
QY	389	TCGAGATGTTTCACTCAGAGATGACAGTTTCTGTATAGACACTTACATCAGAGATCCCA	448
DB	1786	TGAAACACATATCTTGAATGTAATGATGCTTCTTATGACACTTACATAAAGATCCCA	1727
QY	449	AGAAAGGGAAATTTTATTTAATGCAATTTGAACACCATGCCCTATGTTAAGAAAAAGCAG	508
DB	1726	AGAAAGGGAATTTCTTTCATGCAATTTGAACAGATGCTTGTGTGTCAGAGAGGCGAG	1667
QY	509	ATTGGGCTTGGATGATAGCAGATAGAAAATCTACTTTTGGGGAAGAGTGGTGGCTT	568
DB	1666	ACTGGGCTTGGCTGGATTTGGGGAACAAAGAGGCTACCTATGCTGAACGTTGTAGCCT	1607
QY	569	TTGCTGCTGTAGAGAGGTTTCTTCTCAGGATCTTTTCTGCTATATTTCTGGCTAAGA	628
DB	1606	TTGCTGCTGTAGAGAGGTTTCTTCTTTCGCGTCTTTTCTGCTGATTTCTGGCTCAAGA	1547
QY	629	AGAGAGGTTCTTATGCCAGACTCCTTTTTCATGAACTCATCAGCAGAGATGAAGGAC	688
DB	1546	AACGAGGACTGATGCTGCGCTCACATTTTCTTAATGAATTTATAGCAGATGAGGGTT	1487

QY	689	TTCACTGTGACTTTGCTTGCCTGATGTTCCAATATCTTAGTAAATAAGCCTTCAGAGAAA	748
DB	1486	TACACTGTGATTTGCTTGCCTGATGTTCAAAACACCTGGTACACAAACCATCGGAGGAG	1427
QY	749	GGGTACGGGAGATCATTTGTTGATGCTGTCAAAATTTGACGAGGATTTTAAACAGAGCCT	808
DB	1426	GAGTAAGAGAAATAATTTATCAATGCTGTTCGATAGAAACAGGAGTTCTCTACTGAGCCT	1367
QY	809	TGCCAGTTGGCCTCATTTGGAATGAATTCATTTTGTATGAAACAGTACATTTGATTTGTAG	868
DB	1366	TGCTGTGAAGCTCATTTGGGATGAATTCGACTCTAAATGAAGCAATACATTTGATTTGTGG	1307
QY	869	CTGACAGATTACTTGTGGAATTCGATTTCTCAAAAGGTTTTTTCAGCAGCAAAAATCCTTTTG	928
DB	1306	CAGACAGACTTATGCTGGAATTCGTTTACCAAGTTTTCAGAGTAGAGAACCATTTTG	1247
QY	929	ATTTTATGAAAAACATTTCTTTTAAAGAAAAACAAATTTCTTTTGAGAAAAACGAGTTTCAG	988
DB	1246	ACTTTATGGAGATATTTCTACTGGAAGGAAAGACTAACTTCTTTTGAGAAAGAGTAGGCG	1187
QY	989	AGTATCAGCGTTTTCAGTTTATGGCAGAAACACAGATACGTTCTTACCTTGGATGCGAG	1048
DB	1186	AGTATCAGAGGATGGGAGTGTGATGTCAAGTCCACAGAGAAATTTCTTTTACCTTGGATGCTG	1127
QY	1049	ATTT 1052	
DB	1126	ACTT 1123	

RESULT 10
ABL65414
ID ABL65414 standard; DNA; 2500 BP.
XX
AC ABL65414;
XX
DT 15-MAY-2002 (first entry)
XX
DE Lung cancer related gene sequence SEQ ID NO:3751.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US010838.
XX
PR 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0233333P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.

28-SEP-2000; 2000US-0236028P.
28-SEP-2000; 2000US-0236032P.
28-SEP-2000; 2000US-0236033P.
28-SEP-2000; 2000US-0236034P.
28-SEP-2000; 2000US-0236109P.
28-SEP-2000; 2000US-0236111P.
29-SEP-2000; 2000US-0236842P.
29-SEP-2000; 2000US-0236891P.
02-OCT-2000; 2000US-0237172P.
02-OCT-2000; 2000US-0237173P.
02-OCT-2000; 2000US-0237278P.
02-OCT-2000; 2000US-0237294P.
02-OCT-2000; 2000US-0237295P.
02-OCT-2000; 2000US-0237316P.
02-OCT-2000; 2000US-0237425P.
03-OCT-2000; 2000US-0237598P.
03-OCT-2000; 2000US-0237604P.
03-OCT-2000; 2000US-0237606P.
03-OCT-2000; 2000US-0237608P.
01-NOV-2000; 2000US-0244867P.
01-NOV-2000; 2000US-0245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
XX
XX
DR WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX
PS Claim 1; SEQ ID NO 3751; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC tumour
XX
SQ Sequence 2500 BP; 679 A; 522 C; 595 G; 704 T; 0 U; 0 Other;

Query Match 55.1%; Score 580; DB 6; Length 2500;
Best Local Similarity 75.1%; Pred. No. 1.4e-154;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;
QY 89 ATGAAGACCACTCTCTAAGAAAGAGTTCTCGCGGGTTGTTCATCTTTCCAAATCCAGTACC 148
DB 397 AGGATGAGCGCTGCTGAGAGAAACCCCGCGGCTTGTTCATCTTCCCATCGAGTACC 456
QY 149 CTGATATTGGAAATGATATAACAGGCACAGGCTTCCTTCTGGACAGAGAGAGTTG 208
DB 457 ATGATATCTGGCAGATGATATAAGAGGCAGAGGCTTCCTTTTGGACCGCGAGGAGTTG 516
QY 209 ACTTATCAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAGTACTTTCATCT 268
DB 517 ACCTCTCCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTATAT 576
QY 269 CTCACATCTTAGCCTTTTGGCAGCGAGTGATGGAAATTTAAATGAAATTTGGTGAGC 328

Db 577 CCATGTTCTGGCTTTCTTTTTCAGCAAGCGATGCGATAGTAAATGAAAACTTGGTGGAGC 636
QY 329 GCTTTAGTCAGGAGGTGCGAGGTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAATTTCTCA 388
Db 637 GATTTAGCCCAAGAGTTTCAGATTACAGAGCCCGCTGTTTCTATGGCTTTCAAATTTGCCA 696
QY 389 TCGAGAAATGTTCACTCAGAGATGTACAGTTTGTGTGATAGACACTTACATCAGAGATCCCA 448
Db 697 TGGAAACACATACATTCTGAAATGTATAGTCTTCTTATTGACACATTACATAAAGATCCCA 756
QY 449 AGAAAAGGGAATTTTATTAAATGCAATGCAACCATGCCCCATGTTTAAAGAAAAGAGCAG 508
Db 757 AAGAAAAGGGAATTTTCTCTCAATGCCATTGAAACGATGCTTGTCTCAAGAGAGAGGAG 816
QY 509 ATTGGGCTTTCGATGATAGCAGATAGAAAATCTACTTTTGGGGAAGAGTGGTGGCT 568
Db 817 ACTGGGCTTTCGCTGATTTGGGACAAAGAGGCTACTATGGTGAACGTGTTGTAGCT 876
QY 569 TTGCTGCTGTAGAAGGAGTTTCTTCTCAGGATCTTTTGTGCTGTATATTTGGCTTAAAGA 628
Db 877 TTGCTGCTGAGTGAAGGCAATTTCTTTCCGGTCTCTTTGCGTGCATATTTCTGGCTCAAGA 936
QY 629 AGAGAGTCTTATGCCAGGACTCACTTTTCCAAATGAACCTCATCAGCAGAGATGAAGAC 688
Db 937 AAGCAGGACTGATGCCCTGGCTCACAATTTCTAATGAACCTTATTAGCAGAGATGAGGTT 996
QY 689 TTCACGTGCACTTTGGCTGCTGATGTTCCAAATCTTAGTAAATAAGCCTTCAGAGAGAA 748
Db 997 TACACTGTGATTTTGTCTGCTGATGTTCAAAACACCTGGTACAAACCATCGGAGGAGA 1056
QY 749 GGGTCAGGAGATCAATTTGTGATGCTCTCAAAATTTGAGCAGGAGTTTAAACAGAGCT 808
Db 1057 GAGTAAGAGAAATAATATCAATGCTCTTCGGATAGAACAGAGTTCTCTCAGGCT 1116
QY 809 TGCAGTGGCCCTCATTTGGAATGAATTTGATGAAACAGTACATTTGAGTTTGTAG 868
Db 1117 TGCCTGTGAAGCTCAATTTGGGATGAATTCGACTCTTAATGAAGCAATACATTTGAGTTTGTGG 1176
QY 869 CTGACAGATTACTTGTGGAACCTTGGATCTCTCAAGTCTTTTTCAGCAGCAAAATCTTTTG 928
Db 1177 CAGACAGACTTATGCTGGAACCTGGGTTTATGAGAGGTTTTCAGAGTAGAGAACCCATTTG 1236
QY 929 ATTTTATGGAACCAATTTCTTTAGAGGAAAAACAAATTTCTTTGAGAAAACGAGTTTCAG 988
Db 1237 ACTTTATGAGATATTTCTGGAAGAAAGACTAATCTTTTGGAGAGAGAGTAGGCG 1296
QY 989 AGTATCAGCGTTTTCAGTATGAGCAAAACCAAGATTAAGCTTTTCACTTGGATGAG 1048
Db 1297 AGTATCAGAGGATGGGAGTGTCAAGTCCAAACAGAGAAATCTTTTACCTTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360
RESULT 11
ABL66517
ID ABL66517 standard; DNA; 2500 BP.
XX
XX ABL66517;
XX AC
XX AC
DT 15-MAY-2002 (first entry)
XX
DE Lung cancer related gene sequence SEQ ID NO:4854.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.

XX	13-DEC-2001.	CC	as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
XX		CC	prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
XX		CC	cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
XX		CC	cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms
XX		CC	tumour
XX		XX	
PR	30-MAY-2001; 2001WO-US010838.	SQ	Sequence 2500 BP; 679 A; 522 C; 595 G; 704 T; 0 U; 0 Other;
PR	05-JUN-2000; 2000US-0209473P.		Query Match 55.1%; Score 580; DB 6; Length 2500;
PR	05-JUN-2000; 2000US-0209531P.		Best Local Similarity 75.1%; Pred. No. 1.4e-154;
PR	18-SEP-2000; 2000US-0233133P.		Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;
PR	18-SEP-2000; 2000US-0233617P.		
PR	20-SEP-2000; 2000US-0234009P.		
PR	20-SEP-2000; 2000US-0234034P.		
PR	20-SEP-2000; 2000US-0234052P.		
PR	22-SEP-2000; 2000US-0234509P.		
PR	22-SEP-2000; 2000US-0234567P.		
PR	25-SEP-2000; 2000US-0234923P.		
PR	25-SEP-2000; 2000US-0235134P.		
PR	25-SEP-2000; 2000US-0235280P.		
PR	26-SEP-2000; 2000US-0235637P.		
PR	26-SEP-2000; 2000US-0235638P.		
PR	27-SEP-2000; 2000US-0235711P.		
PR	27-SEP-2000; 2000US-0235720P.		
PR	27-SEP-2000; 2000US-0235840P.		
PR	27-SEP-2000; 2000US-0235863P.		
PR	28-SEP-2000; 2000US-0236028P.		
PR	28-SEP-2000; 2000US-0236032P.		
PR	28-SEP-2000; 2000US-0236033P.		
PR	28-SEP-2000; 2000US-0236034P.		
PR	28-SEP-2000; 2000US-0236109P.		
PR	28-SEP-2000; 2000US-0236111P.		
PR	29-SEP-2000; 2000US-0236842P.		
PR	29-SEP-2000; 2000US-0236891P.		
PR	02-OCT-2000; 2000US-0237172P.		
PR	02-OCT-2000; 2000US-0237173P.		
PR	02-OCT-2000; 2000US-0237278P.		
PR	02-OCT-2000; 2000US-0237294P.		
PR	02-OCT-2000; 2000US-0237295P.		
PR	02-OCT-2000; 2000US-0237316P.		
PR	03-OCT-2000; 2000US-0237425P.		
PR	03-OCT-2000; 2000US-0237598P.		
PR	03-OCT-2000; 2000US-0237604P.		
PR	03-OCT-2000; 2000US-0237606P.		
PR	03-OCT-2000; 2000US-0237608P.		
PR	01-NOV-2000; 2000US-0244867P.		
PR	01-NOV-2000; 2000US-0245084P.		
XX		XX	
PA	(AVAL-) AVALON PHARM.		
XX		XX	
PI	Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;		
PI	Soppet DR, Weaver Z;		
XX		XX	
XX	WPI; 2002-188264/24.		
XX		XX	
PT	Screening for anti-neoplastic agent involves exposing cells to a chemical		
PT	agent to be tested for anti-neoplastic activity, and determining a change		
PT	in expression of a gene of a signature gene set.		
XX		XX	
PS	Claim 1; SEQ ID NO 4854; 44pp; English.		
XX		XX	
CC	The present invention describes a method (M1) for screening for an anti-		
CC	neoplastic agent. The method involves exposing cells to a chemical agent		
CC	to be tested for anti-neoplastic activity, determining a change in		
CC	expression of at least one gene (I) of a signature gene set, where (I)		
CC	comprises a sequence (S) selected from 8447 sequences (given in ABL61664		
CC	to ABL70110), or is at least 95% identical to (S), where a change in		
CC	expression is indicative of anti-neoplastic activity. (II) has cytostatic		
CC	activity and can be used in gene therapy. M1 can be used for screening an		
CC	anti-neoplastic agent, and can be used for producing a product which is		
CC	the data collected with respect to the anti-neoplastic agent as a result		
CC	of M1, and the data is sufficient to convey the chemical structure and/or		
CC	properties of the agent. M1 can be used in the treatment of cancer such		

Db 1297 AGTATCAGAGGATGGAGTGATGTCAGTCCACAGAGAATTTCTTTTACCTTGGATGCTG 1356
Qy 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 12
ABL65859
ID ABL65859 standard; DNA; 2500 BP.
XX
AC ABL65859;
XX
DT 15-MAY-2002 (first entry)
XX
DE Lung cancer related gene sequence SEQ ID NO:4196.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; db.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US010838.
XX
PR 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0233133P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235072P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 01-NOV-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.

XX (AVAL-) AVALON PHARM.
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX
PS Claim 1; SEQ ID NO 4196; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC tumour
XX
SQ Sequence 2500 BP; 679 A; 522 C; 595 G; 704 T; 0 U; 0 Other;
Query Match 55.1%; Score 580; DB 6; Length 2500;
Best Local Similarity 75.1%; Pred. No. 1.4e-154;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;
Qy 89 ATGAAGAGCCACTCTTAAGAAAGAGTCTCGCGGTTTGTCATCTTTCCATCCAGTACC 148
Db 397 AGGATGAGCGCGCTGCTGAGAGAAACCCCGCGCTTGTGTCATCTCCCATCGAGTACC 456
Qy 149 CTGATATTTGGAAATGTATAAAACAGGCACAGGCTTCTCTTGACACAGAGAGGTG 208
Db 457 ATGATATCTGCGAGATGATAAGAGGCAGAGGCTTCTTTTGGACCGCCGAGGAGGTG 516
Qy 209 ACTTATCAAGAGTCTCCCTCACTGGAAACAAGCTTAAAGCAGATGAGAAGTACTTCACT 268
Db 517 ACCTCTCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTATAT 576
Qy 269 CTCACATCTTAGCTTTTGGAGCCAGTGATGGAATGTAAATGAAATTTGGTGGAGC 328
Db 577 CCATGTTCTGGCTTCTTTTGCAGCAAGCGATGCGATAGTAATGAAATCTTGGTGGAGC 636
Qy 329 GCTTTAGTCAGGAGGTGCGAGGTTCCAGAGGCTCGCTGTTCTATGGCTTCAAAATCTCA 388
Db 637 GATTTAGCCAAAGAGTTCAGATTACAGAGCCCGCTGTTCTATGGCTTCAAAATGCA 696
Qy 389 TCGAGATGTTCACTCAGAGATGTACAGTTTGTGTAGATAGACACTTACATCAGATCCCA 448
Db 697 TGGAAACATACATTCTGAAATGTATAGTCTTCTTATGACACTTACATAAAGATCCA 756
Qy 449 AGAAAAGGGAATTTTATTAATGCAATGGAACCATGCCCTATGTTAAAGAAAAAGCAG 508
Db 757 AAGAAAGGGAATTTCTCTCAATGCCATTGAAACGATGCGCTTGTGTCAAGAAGAGGCAG 816
Qy 509 ATTGGGCTTGGATGATAGATAGAAATCTACTTTTGGGGAGAGTGGTGGCT 568
Db 817 ACTGGGCTTGGCGCTGGATTGGGGACAAAGAGGCTACTATGGTGAACGTGTTAGCT 876
Qy 569 TTGCTGCTAGAGAGGAGTTTCTTCTCAGGATCTTTTGTGCTATATTTCTGGCTAAAGA 628
Db 877 TTGCTGCGAGTGGAGGCAATTTCTTTTCCGGTCTTTTGTGCTCGATATTTCTGGCTCA 936

QY 569 TTGCTGCTGTAGAAGAGTTTCTTCTCAGGATCTTTTCTGCTATATTTCTGCTAAAGA 628
Db 877 TTGCTGCACTGGAAGGCAATTTCTTTTCGGTCTTTTTCGGTCTGATATTTCTGCTCAAGA 936
QY 629 AGAGAGGCTTTATGCCAGACTCATTCTTCCAAATGCACTATCAGCAGATGAAGAC 688
Db 937 AACGAGGACTGATGCTGCGCTCACATTTTCTAATGAATTTATAGCAGAGATGAGGGTT 996
QY 689 TTCCTGTGACTTTGCTGCTGATGTTCCAAATCTAGTAAATAGCCTTCAGAGAAA 748
Db 997 TACACTGTGATTTGCTGCTGATGTTCAACACTGTTACAAACCATCGAGGAGA 1056
QY 749 GGGTCAGGAGATCAATGTTGATGCTGCTCAAAATTTGAGCAGGAGTTTAAACAGAGCCT 808
Db 1057 GAGTAAGAGAAATATATCAATGCTGTTCCGATAGAACAGGAGTTTCTCCTCAGTGGCCT 1116
QY 809 TGCAGTTGGCTCATTTGAATGAATGCAATTTTGTATGAAACAGTACATGATTTGTAG 868
Db 1117 TGCCTGTGAAGCTCATTTGGGATGAATTCACCTTAATGAAGCAATACATTTGATTTGTGG 1176
QY 869 CTGACAGATTACTGTGGAACCTTGGATTCTCAAGGTTTTCAGGCAGAGAAATCCTTTTG 928
Db 1177 CAGACAGACTTATGCTGGAACCTGGGTTTATAGCAAGGTTTTCAGATAGAGAACCCATTTG 1236
QY 929 ATTTATGAAACCAATTTCTTTAGAGGAGAAACAAATTTCTTTGAGAAACGAGTTTTCAG 988
Db 1237 ACTTTATGAGAAATATTTCACTGGAAGGAAAGCACTAACTTTCTTTGAGAGAGATGAGCG 1296
QY 989 AGTATCAGCGTTTTCAGTTATGGCAGAGAAACACAGATAACGTTCTTACCTTGGATGCG 1048
Db 1297 AGTATCAGAGGATGGGAGTGATGTCAAGTCCAAACAGAGAAATCTTTTACCTTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360
RESULT 14
ADI32044
ID ADI32044 standard; cDNA; 2500 BP.
XX
AC ADI32044;
DT 17-JUN-2004 (first entry)
DE Human cDNA #1370.
XX
KW Human; gene; ss; immunological response; immunopathological condition;
KW Crohn's disease; asthma; ulcerative colitis; hypereosinophilia;
KW irritable bowel syndrome; osteoarthritis; rheumatoid arthritis;
KW acute monocytic leukaemia; antinflammatory; antiasthmatic; antiulcer;
KW osteopathic; antiarthritic; antirheumatic; cytostatic.
XX
OS Homo sapiens.
XX
PN US6607879-B1.
XX
PD 19-AUG-2003.
XX
PF 09-FEB-1998; 98US-00023655.
XX
PR 09-FEB-1998; 98US-00023655.
XX
PA (INCY-) INCYTE CORP.
XX
PI Cocks BG, Stuart SG, Seilhamer JU;
XX
DR WPI; 2003-895307/82.
XX
PT A composition comprising a plurality of cDNAs, useful for detecting
PT altered expression of genes in an immunological response or for
PT diagnosing and treating an immunopathology, e.g. Crohn's disease, asthma
PT or osteoarthritis.

XX
PS
XX
CC The invention relates to a composition comprising a plurality of cDNAs
CC for detecting the altered expression of genes in an immunological
CC response. The invention also relates to a method of diagnosing or
CC monitoring the treatment of an immunopathological condition in a sample,
CC comprising obtaining nucleic acids from a sample, contacting the nucleic
CC acids of the sample with an array comprising the plurality of cDNAs under
CC conditions to form one or more hybridisation complexes, detecting the
CC hybridisation complexes and comparing the levels of the detected
CC hybridisation complexes with the level of hybridisation complexes
CC detected in a non-diseased sample, where an altered level of the detected
CC hybridisation complexes correlates with the presence of an
CC immunopathological condition. Also disclosed are an expression profile
CC comprising a microarray and a plurality of detectable complexes and a
CC method for identifying a plurality of polynucleotide probes. The cDNAs
CC are useful as hybridisable array elements in a microarray for monitoring
CC the expression of target polynucleotides. The microarray can be used in
CC the diagnosis of an immunopathology, such as Crohn's disease, asthma,
CC ulcerative colitis, hypereosinophilia, irritable bowel syndrome,
CC osteoarthritis, rheumatoid arthritis or acute monocytic leukaemia, and in
CC identifying agents for the treatment of the diseases. The microarray may
CC also be used in drug discovery and development, toxicological and
CC carcinogenicity studies, forensics or pharmacogenomics. The composition
CC may also be used in purification of a subpopulation of mRNAs, cDNAs or
CC genomic fragments. This sequence represents a human cDNA of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification but was obtained in electronic format directly
CC from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 2500 BP; 679 A; 522 C; 595 G; 704 T; 0 U; 0 Other;
Query Match 55.1%; Score 580; DB 11; Length 2500;
Best Local Similarity 75.1%; Pred. No. 1.4e-154;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;
QY 89 ATGAAGAGCCACTCTTAAGAAAGAGTTCTCGCGGTTTGTTCATCTTTTCCAATCCAGTACC 148
Db 397 AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTTGTTCATCTTTCCCATCCAGTACC 456
QY 149 CTGATATTTGGAAAATGTATAAACAGGCACAGGCTTCTTTCTGGACAGCAGAGAGTTG 208
Db 457 ATGATATCTGGCAGATGTATAAGAAGCGCAGAGGCTTCTTTTGGACCGCCGAGAGTTG 516
QY 209 ACTTATCAAGGATCTCCCTCACTGGAACAAAGCTTAAAGCAGATGAGAGTACTTCATCT 268
Db 517 ACCTCTCCAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTATAT 576
QY 269 CTCACATCTTAGCCTTTTTCAGCCAGTCATGGAATGTAAATGAAATTTGGTGGAGC 328
Db 577 CCCATGTTCTGGCTTCTTTTTCAGCAAGCGATGCAATAGTAAATGAAACCTTGGTGGAGC 636
QY 329 GCTTTAGTCAGGAGGTCAGGTTCCAGAGGCTCGCTTCTTCTATGGCTTTCAAAATTTCTCA 388
Db 637 GATTTAGCCAAGAAGTTTCAGATTACAGAACCCGCTCTTTCTATGGCTTCCAAATTGCCA 696
QY 389 TCAGAAATGTTCACTCAGAGATGTACAGTTTGTGTGATAGACACTTACATCAGAGATCCCA 448
Db 697 TGGAAAAACATACATTTCTGAAATGTATAGTCTTCTTATTGACACTTACATAAAGATCCCA 756
QY 449 AGAAAGGGGAAATTTTATTTAATGCAATTCGAACCATGCCCCATGTTTGAAGAAAAGCAG 508
Db 757 AAGAAAGGGGAAATTTCTCTCAATGCCCATTGAAACGATGCGCTTGTGTCAAGAAAGGCGAG 816
QY 509 ATTGGGCTTGGCATGAGATAGAGATAAGAAATCTACTTTTGGGCAAGAGTGGTGGCT 568
Db 817 ACTGGGCTTGGCTGGATTGGGACAAAGAGGCTACTATGGTGAACGTTGTGTAGGCT 876
QY 569 TTGCTGCTGTAGAAGGAGTTTCTTCTCAGGATCTTTTGTGCTGTATATTTCTGGCTAAAGA 628
Db 877 TTGCTGCACTGGAAGGCAATTTCTTTTTCGGTCTTTTTCGGTCTGATATTTCTGCTCAAGA 936

Db 1357 ACTT 1360

Search completed: October 30, 2005, 00:02:46
Job time : 442.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2005, 22:53:45 ; Search time 2718.5 Seconds
(without alignments)
14744.041 Million cell updates/sec

Title: US-10-698-228-12
Perfect score: 1053
Sequence: 1 atggcgaccgaaagcc.....tcacottgagcagatttt 1053

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsal:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1051.4	99.8	3292	3	CR617553 full-leng
2	1051.4	99.8	4748	3	AL137348 Homo sapi
3	1039.4	98.7	4650	3	BC042948 Homo sapi
4	1009	95.8	3501	3	CR627376 Homo sapi
5	943.4	89.6	1056	9	AY398973 Homo sapi
6	928	88.1	1083	1	AL547501 AL547501
7	828.2	78.7	1008	9	AY398974 Pan trogl
8	791.6	75.2	879	5	BQ441857 AGENCOURT
9	778.6	73.9	1043	9	AY398975 Mus muscu
10	731.2	69.4	892	7	CN163214 952695 MA
11	726.6	69.0	914	7	CF995079 AGENCOURT
12	710.6	67.5	870	7	CO648380 ILLUMIGEN
13	682.8	64.8	804	6	CD566661 AGENCOURT
14	653.8	62.1	1061	5	BM801298 AGENCOURT
15	653.2	62.0	896	5	BUI70979 AGENCOURT
16	614.8	58.4	685	5	BX951374 DKFZp781E
17	578.4	54.9	1571	3	CR603461 full-leng
18	578.4	54.9	1573	3	CR625489 full-leng
19	578.4	54.9	1582	3	CR608076 full-leng
20	578.4	54.9	1588	3	CR602054 full-leng
21	578.4	54.9	1592	3	CR621427 full-leng
22	578.4	54.9	1600	3	CR604378 full-leng
23	578.4	54.9	1605	3	CR590959 full-leng
24	578.4	54.9	1612	3	CR614990 full-leng

25	578.4	54.9	1613	3	CR609838 full-leng
26	578.4	54.9	1623	3	CR618451 full-leng
27	578.4	54.9	1630	3	CR602150 full-leng
28	578.4	54.9	1796	3	CR596700 full-leng
29	570.4	54.2	2088	3	AK088907 Mus muscu
30	560.4	53.2	638	7	CN358022 170005831
31	545.4	51.8	547	7	CN358025 170005326
32	529.2	50.3	836	7	CR416843 CR416843
33	527.2	50.1	608	1	AU137050 AU137050
34	525.6	49.9	540	5	BU429508 UI-HF-BNO
35	525	49.9	997	4	BM468712 AGENCOURT
36	518.8	49.3	522	5	BQ694828 1000971 H
37	513.6	48.8	793	7	CK778582 965745 MA
38	503	47.8	697	6	CA328968 UI-N-FYO-
39	503	47.8	792	5	BU328458 603494277
40	502.2	47.7	923	7	CR580780 CR580780
41	502	47.7	510	7	CR540284 DKFZp459N
42	494.8	47.0	946	7	CN024619 AGENCOURT
43	492.2	46.7	510	5	BX955728 DKFZp781G
44	491.8	46.7	932	5	BUI96941 AGENCOURT
45	489.4	46.5	909	6	CA981614 AGENCOURT

ALIGNMENTS

RESULT 1
CR617553
LOCUS
DEFINITION
full-length cDNA clone CS0DI011YF14 of Placenta Cot 25-normalized
of Homo sapiens (human).
ACCESSION
CR617553.1 GI:50498360
VERSION
HTC; CNSLT_CDNA.
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 3292)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
Genoscope.
REFERENCE
2 (bases 1 to 3292)
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
Location/Qualifiers
source
1. .3292
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI011YF14"
/tissue type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

Query Match 99.8%; Score 1051.4; DB 3; Length 3292;
Best Local Similarity 99.9%; Pred. No. 1.5e-268;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGCGACCCGAAAGCGCGGCGGCTGATCAGATGAGATCATCTTCA 60
|||||
DB 27 ATGGCGACCCGAAAGCGCGGCGGCTGATCAGATGAGATCATCTTCA 86
|||||

QY	61	GACACCAACGAAGTGAATAAAGTCAAAATGAAGAGCCA	CTCCTAAGAAAGAGTTCTCGC	120
Db	87	GACACCAACGAAGTGAATAAAGTCAAAATGAAGAGCCA <th>CTCCTAAGAAAGAGTTCTCGC</th> <td>146</td>	CTCCTAAGAAAGAGTTCTCGC	146
QY	121	CGGTTTGTCATCTTTCCAAATCCAGTACCCCTGATATTTGGAAATGATATAAACAGGCACAG <td>180</td>	180	
Db	147	CGGTTTGTCATCTTTCCAAATCCAGTACCCCTGATATTTGGAAATGATATAAACAGGCACAG <td>206</td>	206	
QY	181	GCCTCCCTTCGACACGACAGAAAGAGTTGACTTATCAAAAGGATCTCCCTCACTCGAACCAAG <td>240</td>	240	
Db	207	GCCTCCCTTCGACACGACAGAAAGAGTTGACTTATCAAAAGGATCTCCCTCACTCGAACCAAG <td>266</td>	266	
QY	241	CTTAAAGCAGATGAGAAAGTACTTCATCTCTCACATCTTTAGCCCTTTTTCGAGCCAGTGAT <td>300</td>	300	
Db	267	CTTAAAGCAGATGAGAAAGTACTTCATCTCTCACATCTTTAGCCCTTTTTCGAGCCAGTGAT <td>326</td>	326	
QY	301	GGAAATGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTTCAGAGGCT <td>360</td>	360	
Db	327	GGAAATGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTTCAGAGGCT <td>386</td>	386	
QY	361	CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAAATGTTCACTCAGAGATGTTACAGTTTG <td>420</td>	420	
Db	387	CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAAATGTTCACTCAGAGATGTTACAGTTTG <td>446</td>	446	
QY	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAAATTTTATTTAATGCAATTGAA <td>480</td>	480	
Db	447	CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAAATTTTATTTAATGCAATTGAA <td>506</td>	506	
QY	481	ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGCCCTTGGATGGATAGCAGATAGAAAA <td>540</td>	540	
Db	507	ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGCCCTTGGATGGATAGCAGATAGAAAA <td>566</td>	566	
QY	541	TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAAGAGATTTTCTTCTCAGGA <td>600</td>	600	
Db	567	TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAAGAGATTTTCTTCTCAGGA <td>626</td>	626	
QY	601	TCTTTTGTCTCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTTCC <td>660</td>	660	
Db	627	TCTTTTGTCTCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTTCC <td>686</td>	686	
QY	661	AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGTATGTTCCAA <td>720</td>	720	
Db	687	AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGTATGTTCCAA <td>746</td>	746	
QY	721	TACTTAGTAAATAGCCCTTCAGAAAGAGGTTCAGGAGATCATGTTGATGCTGTCAAA <td>780</td>	780	
Db	747	TACTTAGTAAATAGCCCTTCAGAAAGAGGTTCAGGAGATCATGTTGATGCTGTCAAA <td>806</td>	806	
QY	781	ATTGAGCAGGAGTTTTTAAACAGAGCCCTTGCAGTTCGGCCTCATTTGGAATGAATTCGATT <td>840</td>	840	
Db	807	ATTGAGCAGGAGTTTTTAAACAGAGCCCTTGCAGTTCGGCCTCATTTGGAATGAATTCGATT <td>866</td>	866	
QY	841	TTGATGAAAACAGTACATGTAGTTTGTAGCTGACAGATTAATTTGGAACCTTGATTCCTCA <td>900</td>	900	
Db	867	TTGATGAAAACAGTACATGTAGTTTGTAGCTGACAGATTAATTTGGAACCTTGATTCCTCA <td>926</td>	926	
QY	901	AAGGTTTTTTCAGGCAGAAAAATCCCTTTTGAATTTATGGAACAACTTTCTTTAGAGGAAA <td>960</td>	960	
Db	927	AAGGTTTTTTCAGGCAGAAAAATCCCTTTTGAATTTATGGAACAACTTTCTTTAGAGGAAA <td>986</td>	986	
QY	961	ACAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC <td>1020</td>	1020	
Db	987	ACAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC <td>1046</td>	1046	
QY	1021	ACAGATAACGCTTTCACCTTTGGATGCGAGATTTT <td>1053</td>	1053	
Db	1047	ACAGATAACGCTTTCACCTTTGGATGCGAGATTTT <td>1079</td>	1079	
RESULT 2				
HSM802033		4748 bp	mRNA	linear
LOCUS	HSM802033	4748 bp	mRNA	linear
				HTC 22-SEP-2004

DEFINITION	Homo sapiens mRNA; cDNA DKFp761E1312 (from clone DKFp761E1312).
ACCESSION	AL137348
VERSION	AL137348.1
KEYWORDS	GI:6807859
SOURCE	HTC.
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens
AUTHORS	Bukaryova; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 4748) Ansoorge.W., Krieger,S., Regiert,T., Rittmuller,C., Schwager,B., Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.
CONSTRMT	The German CDNA Consortium
TITLE	Direct Submission
JOURNAL	Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
COMMENT	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by EMBL (European Molecular Biology Laboratories, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFp761E1312) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering: http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFp761E1312 Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.
FEATURES	Location/Qualifiers
source	1..4748
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="RZPD:DKFp761E1312"
	/db_xref="taxon:9606"
	/clone="DKFp761E1312"
	/tissue_type="anygdala"
	/clone_lib="761 (synonym: hamy2). Vector pSport1; host DH10B; sites NotI + SalI"
	/dev_stage="adult"
	/notes="ribonucleotide reductase"
gene	1..4748
	/gene="DKFp761E1312"
CDS	47..1102
	/gene="DKFp761E1312"
	/codon_start=1
	/product="hypothetical protein"
	/protein_id="CA870703.2"
	/db_xref="GI:52545612"
	/db_xref="GOA:Q9NTD8"
	/db_xref="UniProt/TREMBL:Q9NTD8"
	/translation="MGDKPRPAAAGVDQDERSSDNTSEISKSNEPRLLRKSRREVI FTIYQPDIMWKYKQAQSPWTAPEVDLSKDLPHWKKADKEYFISHILAFPAASDGI VNIENVERESQVPEPACRFYGFQILLIENHSEMYSLIIDVIRDPKPKRELFNAIE THPYVKKADMALRWIADRKSTFGERVAFVAVEGFFSGSFAALFWLKKRGMLPGILT FNNELISDEGLHDLFACMLQYLVNKPSEERVREILIVDAKIEQFLEALPVLGILG MNCILMKQYIEFVADRLRLVELGFSKVFOAENPFDFMENISLEGKTNFFEKRVSEYQRF AVMAETTDNVFTLDADF"
ORIGIN	
Query Match	99.8%; Score 1051.4; DB 3; Length 4748;
Best Local Similarity	99.9%; Pred. No. 1.6e-268;
Matches 1052; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
Qy	1 ATGGGGCAGCCGGAAGCGCGGAGCGCGCGCTGGATCAGATCAGAGATCATCTTCA 60
Db	47 ATGGGGCAGCCGGAAGCGCGGAGCGCGCGCTGGATCAGATCAGAGATCATCTTCA 106
Qy	61 GACACCAAGCAAGTGAATAAAGTCAAAATGAAGGCCACTCCTCAAGAAAGAGTTCCTCG 120
Db	107 GACACCAAGCAAGTGAATAAAGTCAAAATGAAGGCCACTCCTCAAGAAAGAGTTCCTCG 166
Qy	121 CGGTTTGTCATCTTTCAATCCAGTACCCCTGATATTTGGAAATATATAACAGGCACAG 180
Db	167 CGGTTTGTCATCTTTCAATCCAGTACCCCTGATATTTGGAAATATATAACAGGCACAG 226

Db 72 ATGGCGCAGCCCGAAGGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 131
QY 61 GACACCAAGCAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAGAGTGTCTCGC 120
Db 132 GACACCAAGCAAGTGAATAAGTCAATGAAGAG-CACTCTTAAGAAGAGTGTCTCGC 190
QY 121 CGGTTTGTTCATCTTCCATCCAGTACCCTGATATTTGGAAATGTATAAACAGGCACAG 180
Db 191 CGGTTTGTTCATCTTCCATCCAGTACCCTGATATTTGGAAATGTATAAACAGGCACAG 250
QY 181 GCTTCCTTCTGACAGCAGAGAGGTGTGATTTCAAGAGATCTCCTCACTGGAAACAG 240
Db 251 GCTTCCTTCTGACAGCAGAGAGGTGTGATTTCAAGAGATCTCCTCACTGGAAACAG 310
QY 241 CTTAAGCAGATGAGAAGTACTTCTCTCATCTTACATCTTAGCCCTTTTTCAGCCAGTGAT 300
Db 311 CTTAAGCAGATGAGAAGTACTTCTCTCATCTTACATCTTAGCCCTTTTTCAGCCAGTGAT 370
QY 301 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
Db 371 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 430
QY 361 CGCTGTTCTATGGCTTCAAAATTTCTCATGAGATGTTCACTCAGAGATGTACAGTTTG 420
Db 431 CGCTGTTCTATGGCTTCAAAATTTCTCATGAGATGTTCACTCAGAGATGTACAGTTTG 490
QY 421 CTGATAGACATTCATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 480
Db 491 CTGATAGACATTCATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 550
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCTTCGGATGATAGATAGAAAA 540
Db 551 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCTTCGGATGATAGATAGAAAA 610
QY 541 TCTATTTTGGGAAAGAGTGTGGCTTTGCTGTGTAGAGAGTGTCTTCTTCAGGA 600
Db 611 TCTATTTTGGGAAAGAGTGTGGCTTTGCTGTGTAGAGAGTGTCTTCTTCAGGA 670
QY 601 TCTTTTGTGCTATATTCTGGCTTAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 660
Db 671 TCTTTTGTGCTATATTCTGGCTTAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 730
QY 661 AATGAACATCATCAGCAGATGAAGACTTCTACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 731 AATGAACATCATCAGCAGATGAAGACTTCTACTGTGACTTTGCTTGCCTGATGTTCCAA 790
QY 721 TACTTAGTAAATPAGCTTTCAGAGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAAA 780
Db 791 TACTTAGTAAATPAGCTTTCAGAGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAAA 850
QY 781 ATTGAGCAGGAGTTTAAACAGAGCTTGCAGATTTGGCTCATTTGGAATGAAATTGCAAT 840
Db 851 ATTGAGCAGGAGTTTAAACAGAGCTTGCAGATTTGGCTCATTTGGAATGAAATTGCAAT 910
QY 841 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTAATCTTGGAATCTTGATTTCTCA 900
Db 911 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTAATCTTGGAATCTTGATTTCTCA 970
QY 901 AAGGTTTTTTCAGGCGAAAAATCTTTTGAATTTTATGAAAAACATTTCTTTAGAGGAAAA 960
Db 971 AAGGTTTTTTCAGGCGAAAAATCTTTTGAATTTTATGAAAAACATTTCTTTAGAGGAAAA 1030
QY 961 ACAATTTCTTTGAGAAACAGTTTTCAGAGTATCAGCGTTTTCAGGTTATGCGAGAAACC 1020
Db 1031 ACAATTTCTTTGAGAAACAGTTTTCAGAGTATCAGCGTTTTCAGGTTATGCGAGAAACC 1090
QY 1021 ACAGATAAGCTCTTACCTTTGGATGCGAGATTTT 1053
Db 1091 ACAGATAAGCTCTTACCTTTGGATGCGAGATTTT 1123

LOCUS CR627376 3501 bp mRNA linear HTC 22-SEP-2004
DEFINITION Homo sapiens mRNA; cDNA DKFZp686M05248 (from clone DKFZp686M05248).
ACCESSION CR627376
VERSION CR627376.1 GI:50949847
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 3501)
AUTHORS Ottenwaelder,B., Obermaier,B., Deutschenbaier,S., Schaiipp,A., Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.
CONSTRM The German cDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by Medigenomix (Martinsried/Germany) within the cDNA sequencing consortium of the German Genome Project.
This clone (DKFZp686M05248) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp686M05248
Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.
FEATURES
Location/Qualifiers
1..3501
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686M05248"
/tissue_type="salivary gland"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
/note="ribonucleotide reductase M2 B (TP53 inducible)"
gene 1..3501
/gene="DKFZp686M05248"
332..1225
/gene="DKFZp686M05248"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH10473.1"
/db_xref="GI:50949848"
/translation="MYKQAOAFTAEVDLSKDLPHNKLKADKVFISHILAFFAA SDGIVNENLVERFSQVQVPEARCFYQILINENHSEMSLLDITIRDPKRREFLF NATETMPYKKKADWALRWIADKSTFGERVWPAFAVEGVFSGFAAIFWLKRGFLM PGLTFSNELISRDEGLHCDPACLMFOYLVNKPPEERVREIIVDAVKIQEFLPTALPV GLIGMNCILMKQXIEFVADRLLVDELGFSKVFQAEENPDFMENISLEKTNFPERKVS EYORFAVMAETTDNVFTLDADF"

Query Match 95.8%; Score 1009; DB 3; Length 3501;
Best Local Similarity 98.5%; Pred. No. 2.9e-257;
Matches 1018; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 21 GGAAGCGCGCGGCTGGATGAGATGAGAGATCATCTTCAGACACCAACGAAAGTGAAT 80
Db 190 GGACGAGCTTGGAGGCTCAGGCCGGGAGATCATCTTCAGACACCAACGAAAGTGAAT 249
QY 81 AAGTCAATGAAGAGCACTCTTAAGAAAGAGTTCTCGCCGGTTTGTTCATCTTCCAAAT 140
Db 250 AAGTCAATGAAGAGCACTCTCTAAGAAAGAGTTCTCGCCGGTTTGTTCATCTTCCAAAT 309
QY 141 CCAGTACCTCGATATTTGGAAATGTATTAACAGGCACAGGCTTCTCTCGACAGAGA 200
Db 310 CCAGTACCTCGATATTTGGAAATGTATTAACAGGCACAGGCTTCTCTCGACAGAGA 369
QY 201 AGAGGTTGACTTATCAAGGATCTCCCTCACTGGAACAAAGCTTAAAGCAGATGAGAAGTA 260
Db 370 AGAGGTTGACTTATCAAGGATCTCCCTCACTGGAACAAAGCTTAAAGCAGATGAGAAGTA 429

```
QY 261 CTTTCATCTCTCATCTTACGCTTTTTCAGCCAGTGTGAGGATGTAATGTAATGAAATTT 320
Db 430 CTTTCATCTCTCATCTTACGCTTTTTCAGCCAGTGTGAGGATGTAATGTAATGAAATTT 489
QY 321 GGTGGAGCCCTTTAGTCAGAGGTGAGGTTCCAGAGGTCGCTGTTCTATGCGCTTCA 380
Db 490 GGTGGAGCCCTTTAGTCAGAGGTGAGGTTCCAGAGGTCGCTGTTCTATGCGCTTCA 549
QY 381 AATTCTCATCGAGATGTTCACTCAGAGATGTACAGTTTGTGATAGACACTTACATCAG 440
Db 550 AATTCTCATCGAGATGTTCACTCAGAGATGTACAGTTTGTGATAGACACTTACATCAG 609
QY 441 AGATCCCAAGAAAGGAATTTTATTTAATGCAATTTGAACCATGCCCTATGTTAAGAA 500
Db 610 AGATCCCAAGAAAGGAATTTTATTTAATGCAATTTGAACCATGCCCTATGTTAAGAA 669
QY 501 AAAAGCAGATTGGGCTTCGATGATAGCAGATAGAAAATCTACTTTTGGGAAAGAGT 560
Db 670 AAAAGCAGATTGGGCTTCGATGATAGCAGATAGAAAATCTACTTTTGGGAAAGAGT 729
QY 561 GGTGGCCCTTGTCTGTAGAGGAGTTTCTTCTCAGGATCTTTTGTGCTATATCTCG 620
Db 730 GGTGGCCCTTGTCTGTAGAGGAGTTTCTTCTCAGGATCTTTTGTGCTATATCTCG 789
QY 621 GCTAAAGAGAGAGGTTCTTATCCAGGACTCACTTTTCCAAATGAACTCATCAGCAGAG 680
Db 790 GCTAAAGAGAGAGGTTCTTATCCAGGACTCACTTTTCCAAATGAACTCATCAGCAGAG 849
QY 681 TGAAGGACTTCACTGTGACTTTGCTTGTGCTGATGTTTCCAAATCTTAGTAAATAAGCCTTC 740
Db 850 TGAAGGACTTCACTGTGACTTTGCTTGTGCTGATGTTTCCAAATCTTAGTAAATAAGCCTTC 909
QY 741 AGAAGAAAGGTCAGGAGATCAITGTTGATGCTGTCAAAATGACAGAGGTTTTTAAC 800
Db 910 AGAAGAAAGGTCAGGAGATCAITGTTGATGCTGTCAAAATGACAGAGGTTTTTAAC 969
QY 801 AGAAGCCTTGCAGTGGCTCATTTGGATGAATTCATTTTGTGAAACAGTACATTCGA 860
Db 970 AGAAGCCTTGCAGTGGCTCATTTGGATGAATTCATTTTGTGAAACAGTACATTCGA 1029
QY 861 GTTGTAGCTGACAGATCTTGTGGAATCTGGATCTCAAAAGTGTTCAGGACAGAAA 920
Db 1030 GTTGTAGCTGACAGATCTTGTGGAATCTGGATCTCAAAAGTGTTCAGGACAGAAA 1089
QY 921 TCCCTTTGATTTATGGAACAATTTCTTTAGAGGAAAAACAATTTCTTTGAGAAACG 980
Db 1090 TCCCTTTGATTTATGGAACAATTTCTTTAGAGGAAAAACAATTTCTTTGAGAAACG 1149
QY 981 AGTTTCAGAGTATCAGCGTTTTCAGATGATGAGAGAACCAAGATTAACGCTTACCTT 1040
Db 1150 AGTTTCAGAGTATCAGCGTTTTCAGATGATGAGAGAACCAAGATTAACGCTTACCTT 1209
QY 1041 GGATGCAATTTT 1053
Db 1210 GGATGCAATTTT 1222

RESULT 5
LOCUS AY398973
DEFINITION Homo sapiens HCM0069 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY398973
VERSION AY398973.1 GI:39754962
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1056)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 1056)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
Location/Qualifiers
source
1..1056
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>1056
/locus_tag="HCM0069"
ORIGIN
Query Match 89.6%; Score 943.4; DB 9; Length 1056;
Best Local Similarity 89.6%; Pred. No. 6.2e-240;
Matches 944; Conservative 0; Mismatches 109; Indels 0; Gaps 0;
QY 1 ATGGGCGACCCGAAAGCCGGAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 ATGGGCGACCCGAAAGCCGGAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
QY 61 GACACCAACGAAAGTGAATTAAGTCAATTAAGAGAGCCATCTCTTAAGAAAGAGTTCGCG 120
Db 61 GACACCAACGAAAGTGAATTAAGTCAATTAAGAGAGCCATCTCTTAAGAAAGAGTTCGCG 120
QY 121 CGGTTTGTCTATCTTCCATCCAGTACCTGATATTTGGAAAATGTATAACAGGCACAG 180
Db 121 CGGTTTGTCTATCTTCCATCCAGTACCTGATATTTGGAAAATGTATAACAGGCACAG 180
QY 181 GCTTCCTCTTGGACAGCAGAGAGGTTGACTTATCAAGGATCTCCCTCAGTGAACCAAG 240
Db 181 GCTTCCTCTTGGACAGCAGAGAGGTTGACTTATCAAGGATCTCCCTCAGTGAACCAAG 240
QY 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTGAT 300
Db 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTGAT 300
QY 301 GGAAATTTAAATGAAAATTTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 360
Db 301 GGAAATTTAAATGAAAATTTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 360
QY 361 CGCTGTTTCTATGGCTTCAAAATTCATCAGAGATGTTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGGCTTCAAAATTCATCAGAGATGTTTCACTCAGAGATGTACAGTTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTATTTAATGCAATTCGA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTATTTAATGCAATTCGA 480
QY 481 ACCATGCCCTATGTTAAGAAAAGAGATTTGGCCCTTCGATGATGATGATGATGATGAT 540
Db 481 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 540
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTGTCTGTAGAGAGGTTTCTTCTCAGGA 600
Db 541 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 600
QY 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAAAGAGAGGTTCTTATGCCAGGACTCACATTTTCC 660
Db 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAAAGAGAGGTTCTTATGCCAGGACTCACATTTTCC 660
```

```
TITLE
JOURNAL PUBMED
REFERENCE 14671302
AUTHORS
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
Location/Qualifiers
source
1..1056
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>1056
/locus_tag="HCM0069"
ORIGIN
Query Match 89.6%; Score 943.4; DB 9; Length 1056;
Best Local Similarity 89.6%; Pred. No. 6.2e-240;
Matches 944; Conservative 0; Mismatches 109; Indels 0; Gaps 0;
QY 1 ATGGGCGACCCGAAAGCCGGAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 ATGGGCGACCCGAAAGCCGGAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
QY 61 GACACCAACGAAAGTGAATTAAGTCAATTAAGAGAGCCATCTCTTAAGAAAGAGTTCGCG 120
Db 61 GACACCAACGAAAGTGAATTAAGTCAATTAAGAGAGCCATCTCTTAAGAAAGAGTTCGCG 120
QY 121 CGGTTTGTCTATCTTCCATCCAGTACCTGATATTTGGAAAATGTATAACAGGCACAG 180
Db 121 CGGTTTGTCTATCTTCCATCCAGTACCTGATATTTGGAAAATGTATAACAGGCACAG 180
QY 181 GCTTCCTCTTGGACAGCAGAGAGGTTGACTTATCAAGGATCTCCCTCAGTGAACCAAG 240
Db 181 GCTTCCTCTTGGACAGCAGAGAGGTTGACTTATCAAGGATCTCCCTCAGTGAACCAAG 240
QY 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTGAT 300
Db 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTGAT 300
QY 301 GGAAATTTAAATGAAAATTTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 360
Db 301 GGAAATTTAAATGAAAATTTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 360
QY 361 CGCTGTTTCTATGGCTTCAAAATTCATCAGAGATGTTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGGCTTCAAAATTCATCAGAGATGTTTCACTCAGAGATGTACAGTTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTATTTAATGCAATTCGA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTATTTAATGCAATTCGA 480
QY 481 ACCATGCCCTATGTTAAGAAAAGAGATTTGGCCCTTCGATGATGATGATGATGATGAT 540
Db 481 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 540
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTGTCTGTAGAGAGGTTTCTTCTCAGGA 600
Db 541 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 600
QY 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAAAGAGAGGTTCTTATGCCAGGACTCACATTTTCC 660
Db 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAAAGAGAGGTTCTTATGCCAGGACTCACATTTTCC 660
```

```
QY 661 AATGAACATCATCAGCAGAGATGAAGACATTCACATGTCATTTGCTTGGCCTGATGTTCCAA 720
DB |||||
QY 661 AATGAACATCATCAGCAGAGATGAAGACATTCACATGTCATTTGCTTGGCCTGATGTTCCAA 720
DB |||||
QY 721 TACTTAGTAATAAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGTCATGTCGTCAAA 780
DB |||||
QY 721 TACTTAGTAATAAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGTCATGTCGTCAAA 780
DB |||||
QY 781 ATTGACGAGGAGTTTTTAACAGAGAGCCTTGGCCAGTTTGGCCTCATTTGGAATGAATTCATTT 840
DB |||||
QY 781 ATTGACGAGGAGTTTTTAACAGAGAGCCTTGGCCAGTTTGGCCTCATTTGGAATGAATTCATTT 840
DB |||||
QY 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTTGTGAACTTGGATTCCTCA 900
DB |||||
QY 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTTGTGAACTTGGATTCCTCA 900
DB |||||
QY 901 AAGGTTTTTCAGGCAGAAAATCCTTTTGTATTTTATGAGAAACATTTCTTTAGAGAGAAA 960
DB |||||
QY 901 AAGGTTTTTCAGGCAGAAAATCCTTTTGTATTTTATGAGAAACATTTCTTTAGAGAGAAA 960
DB |||||
QY 961 ACAAAATTTCTTTGAGAAAACAGTTTCAGAGATATCAGCGTTTTCAGTTTATGGCAGAAAAC 1020
DB |||||
QY 961 ACAAAATTTCTTTGAGAAAACAGTTTCAGAGATATCAGCGTTTTCAGTTTATGGCAGAAAAC 1020
DB |||||
QY 1021 ACAGATAACGCTCTTCACTTGGATGCGAGATTTT 1053
DB |||||
QY 1021 ACAGATAACGCTCTTCACTTGGATGCGAGATTTT 1053
DB |||||
```

RESULT 6

```
AL547501
LOCUS 1083 bp mRNA linear EST 25-MAR-2004
DEFINITION AL547501 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1011YF14 5-PRIME, mRNA sequence.
ACCESSION AL547501
VERSION AL547501.3 GI:45747948
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1083)
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:31269332.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with NotI and cloned
into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4436.r
```

For more information about this cluster, see
<http://www.genoscope.cns.fr/cdna?s=CS0D1011DC070P1&c=4436.r>.

FEATURES

```
source
1..1083
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1011YF14"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/issue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized."
```

ORIGIN

Query Match 88.1%; Score 928; DB 1; Length 1083;

```
Best Local Similarity 97.9%; Pred. No. 7.9e-236;
Matches 975; Conservative 6; Mismatches 11; Indels 4; Gaps 4;
QY 1 ATGGGCGACCCGGAAGCGGGAAGCGCGG-GCTGGATCAGGATCAGAGATCATCTTTC 59
DB |||||
QY 27 ATGGGCGACCCGGAAGCGGGAAGCGCGG-GCTGGATCAGGATCAGAGATCATCTTTC 86
DB |||||
QY 60 AGACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTCTAAGAAAGATTTCTCG 119
DB |||||
QY 87 AGACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTCTAAGAAAGATTTCTCG 146
DB |||||
QY 120 CCGGTTTGTCACTCTTCCAAATCAGTACCTCTGATATTTGGAAAAATGTATAAACAGGCACA 179
DB |||||
QY 147 CCGGTTTGTCACTCTTCCAAATCAGTACCTCTGATATTTGGAAAAATGTATAAACAGGCACA 206
DB |||||
QY 180 GCGTTCTCTTCTGGACAGCAGAGAGGTGTACTTTATCAAAAGATCTCCCTCACTGGAACAA 239
DB |||||
QY 207 GCGTTCTCTTCTGGACAGCAGAGAGGTGTACTTTATCAAAAGATCTCCCTCACTGGAACAA 266
DB |||||
QY 240 GCTTAAAGCAGATGAGAAGTACTTCTCATCTCTCACTCTTAGCCCTTTTTCAGCCAGTGA 299
DB |||||
QY 267 GCTTAAAGCAGATGAGAAGTACTTCTCATCTCTCACTCTTAGCCCTTTTTCAGCCAGTGA 326
DB |||||
QY 300 TGAATTTGTAATGAAAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCGAGTTCAGAGGC 359
DB |||||
QY 327 TGAATTTGTAATGAAAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCGAGTTCAGAGGC 386
DB |||||
QY 360 TCGCTCTTCTATGGCTTTCAAAATCTCATCTGAGAAATGTTCACTCAGAGATGTACAGTTT 419
DB |||||
QY 387 TCGCTCTTCTATGGCTTTCAAAATCTCATCTGAGAAATGTTCACTCAGAGATGTACAGTTT 446
DB |||||
QY 420 GCTGTAGACACTTACATCAGAGATCCCAAGAAAAAGGAATTTTATTTAATGCAATTGA 479
DB |||||
QY 447 GCTGTAGACACTTACATCAGAGATCCCAAGAAAAAGGAATTTTATTTAATGCAATTGA 506
DB |||||
QY 480 AACCATGCCCTATGTTAAGAAAAACAGATTTGGGCTTTCGAGTGGATAGCAGATAGAAA 539
DB |||||
QY 507 AACCATGCCCTATGTTAAGAAAAACAGATTTGGGCTTTCGAGTGGATAGCAGATAGAAA 566
DB |||||
QY 540 ATCTACTTTTGGGAAAGAGTGTGGCTTTCGCTGTAGAGAGGATTTTCTTCTCAGG 599
DB |||||
QY 567 ATCTACTTTTGGGAAAGAGTGTGGCTTTCGCTGTAGAGAGGATTTTCTTCTCAGG 626
DB |||||
QY 600 ATCTTTTGTCTGTATATTTCTGGCTTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTC 659
DB |||||
QY 627 ATCTTTTGTCTGTATATTTCTGGCTTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTC 686
DB |||||
QY 660 CAATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTCTTGCCTGATGTTCCA 719
DB |||||
QY 687 CAATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTCTTGCCTGATGTTCCA 746
DB |||||
QY 720 ATCTTAGTAAATAAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGATGCTGTCAAA 779
DB |||||
QY 747 ATCTTAGTAAATAAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGATGCTGTCAAA 806
DB |||||
QY 780 AATTGACGAGGAGTTTTTAACAGAGCCTTGCAGTTTGGCTCATTTGGAATGAATTCGAT 839
DB |||||
QY 807 AATTGACGAGGAGTTTTTAACAGAGCCTTGCAGTTTGGCTCATTTGGAATGAATTCGAT 866
DB |||||
QY 840 TTTGATGAAACAGTACATTCAGTTTGTAGTCGACAGATTAATTTG-TGGAACCTTGGATTC 898
DB |||||
QY 867 TTTGATGAAACAGTACATTCAGTTTGTAGTCGACAGATTAATTTG-TGGAACCTTGGATTC 926
DB |||||
QY 899 CAAAGGTTTTTCAGGCAGAAAATCCTTTTGAATTTATGAAAAACATTTCTTTAGAGAAA 958
DB |||||
QY 927 CAAAGGTTTTTCAGGCAGAAAATCCTTTTGAATTTATGG-AAAMAATTTCTTTAGAGG-A 984
DB |||||
QY 959 AAACAATTTCTTTGAGAAACGAGTTTCAGAGTATC 994
DB |||||
QY 985 AAAMAATTTCTTTKAGAACGAGTTTMRAGWWCAGC 1020
DB |||||
```

RESULT 7


```
/lab host="DH10B (phage-resistant)"
/clone lib="NIH_MGC_92"
/organ="Mus musculus"; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dr primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 75.2%; Score 791.6; DB 5; Length 879;
Best Local Similarity 98.6%; Pred. No. 1.6e-199;
Matches 820; Conservative 0; Mismatches 9; Indels 3; Gaps 2;

QY 1 ATGGGCGACCCGGAAGCGCGAAGCGCGCGGCTGGATCAGATCAGAGATCATCTTCA 60
DB 22 ATGGGCGACCCGGAAGCGCGAAGCGCGCGGCTGGATCAGATCAGAGATCATCTTCA 81

QY 61 GACACCAACGAAAGTCAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
DB 82 GACACCAACGAAAGTCAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 141

QY 121 CGGTTTGTTCATCTTCCAAATCAGTACCTGTATATTTGGAAAATGTATATAACAGGCACAG 180
DB 142 CGGTTTGTTCATCTTCCAAATCAGTACCTGTATATTTGGAAAATGTATATAACAGGCACAG 201

QY 181 GCTTCTCTCTGACAGCAGAGAGGTGACTTATCAAGAGATCTCCCTCACTGGACACAG 240
DB 202 GCTTCTCTCTGACAGCAGAGAGGTGACTTATCAAGAGATCTCCCTCACTGGACACAG 261

QY 241 CTTAAAGCAGATGAGAAGTACTTCTCTCATCTTACATCTTAGCCCTTTTTCGAGCCAGTGAT 300
DB 262 CTTAAAGCAGATGAGAAGTACTTCTCTCATCTCTACATCTTAGCCCTTTTTCGAGCCAGTGAT 321

QY 301 GGAATTGTAATGAAAAATTTGGTGAGCGCTTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
DB 322 GGAATTGTAATGAAAAATTTGGTGAGCGCTTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 381

QY 361 CGCTGTTCTATGGCTTTCAAAATCTCATCGAGATGTTTCACTCAGAGATGTACAGTTTG 420
DB 382 CGCTGTTCTATGGCTTTCAAAATCTCATCGAGATGTTTCACTCAGAGATGTACAGTTTG 441

QY 421 CTGATAGACACTTATACATCAGAGATCCCAAGAAAAAGGAAATTTTATTTAATGCAATTGAA 480
DB 442 CTGATAGACACTTATACATCAGAGATCCCAAGAAAAAGGAAATTTTATTTAATGCAATTGAA 501

QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTCGATGAGATAGCAGATAGAAAA 540
DB 502 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTCGATGAGATAGCAGATAGAAAA 561

QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 600
DB 562 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 621

QY 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGGTTCTTATGCCAGGACTCACTTTTTC 660
DB 622 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGGTTCTTATGCCAGGACTCACTTTTTC 681

QY 661 AATGAACCTCATCAGCAGAGATCAAGAGCTTCACTGTGACTTTTGTCTGCTGATGTTCCAA 720
DB 682 AATGAACCTCATCAGCAGAGATCAAGAGCTTCACTGTGACTTTTGTCTGCTGATGTTCCAA 741

QY 721 TACTTAGTAAATAGCCTTCAGAGAAAGGG--TCAGGGAGATCATTTGTTGATGCTGTCA 778
DB 742 TACTTAGTAAATAGCCTTCAGAGAAAGGGGTCAGGGGAGATCATTTGTTGATGCTGTCA 801

QY 779 AAATT-GAGCAGAGTTTTTAAACAGAGCCTTTGCCAGTTGGCTCACTTCGAA 829
DB 802 AAATTGGAGCAGAGTTTTTAAACAGAGCCTTTGCCAGTTGGCTCAATTGGA 853

RESULT 9
AY398975 1043 bp DNA linear GSS 12-DEC-2003
LOCUS
```

```
Mus musculus HCM0069 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY398975
VERSION AY398975.1 GI:39754964
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1043)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferrieres,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1043)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferrieres,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..1043
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>1043
/locus_tag="HCM0069"
ORIGIN
Query Match 73.9%; Score 778.6; DB 9; Length 1043;
Best Local Similarity 81.3%; Pred. No. 5e-196;
Matches 855; Conservative 0; Mismatches 184; Indels 13; Gaps 1;

QY 1 ATGGGCGACCCGGAAGCGCGAAGCGCGGCTGGATCAGGATGAGATCATCTTCA 60
DB 1 ATGGGCGACCCGGAAGCGCGAAGCGCGGCTGGATGAGGATGAGATCTGTTGCG 60

QY 61 GACACCAACGAAAGTGAATATAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
DB 61 GACACCAACGAAAGTGAATATAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120

QY 121 CGGTTTGTTCATCTTCCAAATCAGTACCTGTATTTGGAAAATGTATAACAGGCACAG 180
DB 121 CGGTTTGTTCATCTTCCAAATCAGTACCTGTATTTGGAAAATGTATAACAGGCACAG 180

QY 181 GCTTCTCTCTGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 240
DB 181 GATCTCTTCTGGACAGCAGAGAGGTTGACTTGTCAAAGGACCTCCCTCACTGGAACAAG 240

QY 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCTCACTTTAGCCCTTTTTCGAGCCAGTGAT 300
DB 241 CTTAAATCCGACGAGAAGTATTTTATCTCCCACTCTTAGCCCTTTTTCGAGCCAGTGAT 300

QY 301 GGAATTGTAATGAAAAATTTGGTGAGCGCTTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
DB 301 GGAATTGTAATGAAAAATTTGGTGAGCGCTTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360

QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGATGTTTCACTCAGAGATGTACAGTTTG 420
DB 361 CGCTGTTTCTATGGCTTTCAAGATTTCTCATCGAGATGTTTCACTCAGAGATGTACAGTTTA 420

QY 421 CTGATAGACACTTATACATCAGAGATCCCAAGAAAAAGGAAATTTTATTTAATGCAATTGAA 480
DB 421 CTAATAGACACTTATACATCAGAGATCCCAAGAAAAAGGAAATTTTATTTAATGCAATTGAA 480
```

QY 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATGTGGCCCTTGCATGGATAGCAGATAGAAAA 540
 Db NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 540
 QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTTGGCTGTGTAGAGAGTTCCTCTCAGGA 600
 Db NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 600
 QY 601 TCTTTTGTCTATATCTTGGCTTAAAGAGAGAGGCTTATGCCAGGAGTCACTTTTTC 660
 Db TCAATTTGTGCAATATCTTGGCTTAAAGAGAGAGGCTTATGCCAGGAGTCACTTTTTC 660
 QY 661 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTTGCCTGTGATGTTCCAA 720
 Db AATGAACATCATCAGCAGGAGTGAAGACTTCACTGTGACTTTTGCCTGTGATGTTCCAG 720
 QY 721 TACTTAGTAATTAAGCTTTCAGAGAAAGGCTCAGGAGATCACTTGTGATGCTGTCAAA 780
 Db TACTTGGTAAACAAGCTTTCAGAGATAGAGTGAAGGAAATCACTGTGATGCTGTTCMA 780
 QY 781 ATTGACGAGGAGTTTTTAACAGAGAGCTTGCAGTGTGGCTCACTTGGAAATGAATTCATT 840
 Db ATCGAG-----CAGAGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 827
 QY 841 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTACTTGTGAACTTGGATTTCTCA 900
 Db TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTACTTGTGAACTTGGATTTCTCA 887
 QY 901 AAGTTTTTTCAGCAGAAAAATCCTTTTGTATTTTATGGAACAACTTTTATGAGGAAAA 960
 Db AAGTTTTTTCAGCAGAAAAATCCTTTTGTATTTTATGGAACAACTTTTATGAGGAAAA 947
 QY 961 ACAATTTTCTTGCAGAAACAGTTTCAGAGATATCAGGTTTTCAGTATGCGAGAAACC 1020
 Db ACAATTTTCTTGCAGAAACAGTTTCAGAGATATCAGGTTTTCAGTATGCGAGAAACC 1007
 QY 1021 ACAGATAACCTCTTCACTTGGATGAGATTT 1052
 Db ACAGATATGCTTCACTTGGATGAGATTT 1039

RESULT 10
 LOCUS CN163214/c 892 bp mRNA linear EST 02-APR-2004
 DEFINITION 952895 MARC 4P1G Sus scrofa cDNA 3', mRNA sequence.
 ACCESSION CN163214
 VERSION CN163214.1 GI:46177644
 KEYWORDS EST.
 SOURCE Sus scrofa (pig)
 ORGANISM Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 1 (bases 1 to 892)
 Smith,T.P.L., Freking,B.A., Ford,J.J., Vallet,J.L., Wise,T.A.,
 Noneman,D.J., Wray,J.E. and Keele,J.W.
 Porcine EST collection using a normalized library constructed from
 embryos representing early developmental stages
 Unpublished (2003)
 JOURNAL Contact: Smith TPL
 COMMENT USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called with phred v0.020425.c and
 trimmed with the aid of the trim_alt option. Vector identified with
 cross_match v0.990329.
 Plate: TMH8060 row: G column: 4
 Seq primer: TAGAAGGCACAGTCGAGG.
 1. .892
 Location/Qualifiers
 /organism="Sus scrofa"
 /mol_type="mRNA"

FEATURES

source

/db xref="taxon:9823"
 /tissue type="pooled"
 /lab host="DH10B"
 /clone lib="MARC 4P1G"
 /note="vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI;
 Library made with combined RNA from day-10, day-13,
 day-15, day-25, and day-30 whole embryos."

ORIGIN

Query Match 69.4%; Score 731.2; DB 7; Length 892;
 Best Local Similarity 94.1%; Pred. No. 2e-183;
 Matches 760; Conservative 0; Mismatches 48; Indels 0; Gaps 0;
 QY 246 AGCAGATGAGAAGTACTTCTCATCTCTAGCCCTTTTTCAGCCAGTGTGGAAT 305
 Db 892 ATCAGATGAGAAGTATTTTATCTCTCACTCTAGCCCTTTTTCAGCCAGTGTGGAAT 833
 QY 306 TGTAAATGAAAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCTCGCTG 365
 Db 832 TGTGAATGAAAACTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCGCTG 773
 QY 366 TTCTATGGCTTTCAAAATCTCATCGAATGTTCACTCAGAGATGTACAGTTGCTGAT 425
 Db 772 TTCTACGGCTTTCAAAATCTCATCGAATGTTCACTCAGAGATGTACAGTTGCTGAT 713
 QY 426 AGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAAACCAT 485
 Db 712 AGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAAACCAT 653
 QY 486 GCCCTATGTTAAAGAAAAAGCAGATTTGGCGCTTCGATGGATAGCAGATAGAAAACTTAC 545
 Db 652 GCCATATGTTAAGAAAAAGCAGATTTGGCGCTTCGATGGATAGCAGATAGAAAACTTAC 593
 QY 546 TTTTGGGAAAGAGTGTGGCTTTTGTCTGTGTAGAGAGGTTTCTCTCAGGATCTTT 605
 Db 592 TTTTGGGAAAGAGTGTGGCTTTTGTCTGTGTAGAGAGGTTTCTCTCAGGATCTTT 533
 QY 606 TGCTGCTATATCTGGCTTAAAGAGAGAGTCTTATGCCAGGACTCAGCTTTTCCAAATGA 665
 Db 532 TGCTGCTATATCTGGCTTAAAGAGAGAGTCTTATGCCAGGACTCAGCTTTTCCAAATGA 473
 QY 666 ACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTCTTGCCTGTGTTTCCAATCTT 725
 Db 472 ACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTCTTGCCTGTGTTTCCAATCTT 413
 QY 726 AGTAAATAAGCCTTTCAGAGAAAGGGTCAGGAGATCATTTGTGATGCTGTCAAAATTGA 785
 Db 412 GSTAAATAAGCCTTTCAGAGCAAGAGTTAGGAGATCATTTGTGATGCTGTGAAATTGA 353
 QY 786 GCAGGAGTTTTTAAACAGAGCCTTGCAGTTGGCTCATTTGGAATGAATTCATTTTGTAT 845
 Db 352 GCAGGAGTTTTTAAACAGAGCCTTGCAGTTGGCTCATTTGGAATGAATTTGTGTTTGTAT 293
 QY 846 GAAACAGTACATTCAGTTTGTAGCTGACAGATTAATTTGTGGAACCTTGCATTTCTCAAGGT 905
 Db 292 GAAACAGTACATTCAGTTTGTAGCTGACAGATTAATTTGTGGAACCTTGCATTTCTCAAGGT 233
 QY 906 TTTTTCAGGCAAAAAATCCTTTTGTATTTTATGAAAAACATTTCTTTAGAGGAAAAACAAA 965
 Db 232 TTTTTCAGGCAAAAAATCCTTTTGTATTTTATGAAAAACATTTCTTTAGAGGAAAAACAAA 173
 QY 966 TTTTCTTTGAGAAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTATTCAGGAGAAACACAGA 1025
 Db 172 TTTTCTTTGAGAAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTATTCAGGAGAAACACAGA 113
 QY 1026 TAAGCTCTTCACTCTGATGCAGATTTT 1053
 Db 112 TAAGCTCTTCACTCTGATGCAGATTTT 85

RESULT 11

CF995079
 LOCUS

CF995079 914 bp mRNA linear EST 25-NOV-2003

```
DEFINITION AGENCOURT 15621478 NIH MGC 147 Homo sapiens cDNA clone
IMAGE:30528806 5', mRNA sequence.
ACCESSION CF995079
VERSION CF995079.1 GI:38511139
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 914)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM615 row: f column: 15
High quality sequence stop: 677.
FEATURES             source
    1..914
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:30528806"
        /tissue_type="Human Placenta"
        /lab_host="DH10B Tona"
        /clone_lib="NIH MGC 147"
        /note="Organ: Placenta; Vector: pBluescriptR; Site_1:
all-XhoI; Site_2: BamH; Oligo-dT primed using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average
insert size 2.3 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: This is
a NIH MGC library."
ORIGIN
Query Match          69.0%; Score 726.6; DB 7; Length 914;
Best Local Similarity 99.0%; Pred. No. 3.4e-182;
Matches 763; Conservative 0; Mismatches 4; Indels 4; Gaps 3;
QY 1 ATGGGCGACCCGGGAAGCGCGGAAGCGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
DB 87 ATGGGCGACCCGGGAAGCGCGGAAGCGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 146
QY 61 GACACCAAGAAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAGAGTTCTCGC 120
DB 147 GACACCAAGAAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAGAGTTCTCGC 206
QY 121 CGTTTGTTCATCTTCCATCCAGTACCTCTGATATTTGGAATGATTAACAGGCACAG 180
DB 207 CGTTTGTTCATCTTCCATCCAGTACCTCTGATATTTGGAATGATTAACAGGCACAG 266
QY 181 GCTTCTCTTGACAGCAGAGAGGTTGACTTATCAAGGATCTCCCTCACTGGAAACAAG 240
DB 267 GCTTCTCTTGACAGCAGAGAGGTCGACTTATCAAGGATCTCCCTCACTGGAAACAAG 326
QY 241 CTTAAAGCAGATGAGAAGTACTTCATCTCTCAATCTTAGCCCTTTTTCGAGCCAGTAT 300
DB 327 CTTAAAGCAGATGAGAAGTACTTCATCTCTCAATCTTAGCCCTTTTTCGAGCCAGTAT 386
QY 301 GGAATTGTAATGAAAATTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGGCTCCAGAGGCT 360
387 GGAATTGTAATGAAAATTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGGCTCCAGAGGCT 446
361 CGCTGTTTCTATGGCTTTCAAATTCATCAGAGATGTTCTCACTCAGAGATGTCAGTTTG 420
447 CGCTGTTTCTATGGCTTTCAAATTCATCAGAGATGTTCTCACTCAGAGATGTCAGTTTG 506
421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTATTAATGCAATGAA 480
507 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTATTAATGCAATGAA 566
481 ACCATGCCCTATGTTAAGAAAAGCAGATGGGCCCTTGGATGGATGATAGAGAGAGAAA 540
567 ACCATGCCCTATGTTAAG-AAAAGCAGATGGGCCCTTGGATGGATGATAGAGAGAAA 624
541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTAGAGAGGTTTCTTCTCAGGA 600
625 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTAGAGAGGTTTCTTCTCAGGA 684
601 TCTTTTGTGCTATATCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 660
685 TCTTTTGTGCTATATCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 744
661 AATGAACATCATCAGCAGAGATGAAGACTTTCACCTGTCATCTTGTGCTGATGTTCAA 720
745 AATGAACATCATCAGCAGAGATGAAGACTTTCACCTGTCATCTTGTGCTGATGTTCAA 804
721 TACTTAGTAATAAGCC-TTCAGAGAAAGGCTC-AGGAGATCATTTGTTG 769
805 TACTTAGTAATAAGCCTTTCAAAAGAAAGGCTTAGGAGATCATTTGTTG 855
RESULT 12
LOCUS CO648380
DEFINITION ILLUMIGEN MQC 41729 Katze MMPB2 Macaca mulatta cDNA clone
IBIUM:25149 5' similar to Bases 5 to 770 highly similar to human
RRM2B (Hs.512592), mRNA sequence.
ACCESSION CO648380.1 GI:50569874
VERSION CO648380
KEYWORDS EST.
SOURCE Macaca mulatta (rhesus monkey)
ORGANISM Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Macaca.
REFERENCE 1 (bases 1 to 870)
AUTHORS Katze, M.G., Thomas, M., Korth, M., Iadonato, S.P. and Magness, C.L.
TITLE Large-scale Rhesus Macaque cDNA Sequencing
JOURNAL Unpublished (2003)
COMMENT Contact: C. Magness
Illumigen Biosciences Inc.
2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
Tel: 2063780400
Fax: 2063780408
Email: cmagness@illumigen.com
Sequenced on 2004.06.02. 691 Q20 bases. Library Preparation: Prof.
Michael Katze Lab at University of Washington DNA Sequencing:
Illumigen Biosciences Inc. For further information, see
http://www.macaque.org
PCR Primers
FORWARD: CCCTCACTAAAGGGAACAAA
BACKWARD: CACTATAGGCGGAATTTGGTA
Insert length: 870 Std Error: 0.00
Plate: CL000348 row: B column: 06
Seq primer: CCCTCACTAAAGGGAACAAA
POLYA=No. Location/Qualifiers
    1..870
        /organism="Macaca mulatta"
        /mol_type="mRNA"
        /strain="Indian"
        /db_xref="taxon:9544"
```



```
/clone="IBIUW:25149"  
/sex="male"  
/tissue_type="blood"  
/cell_type="PBMC"  
/dev_stage="adult"  
/lab_host="Electromax DH10B"  
/clone_lib="Katze MMPB2"  
/note="Vector: pDONR 222; Site  
created from CloneMiner cDNA  
catalog #18249-029"
```

ORIGIN

```
Query Match          67.5%; Score 710.6; DB 7; Length 870;
Best Local Similarity 98.8%; Pred. No. 6.2e-178;
Matches 716; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
```

[illegible]

RESULT 13

Accession	LOCUS	CD656661	804 bp	mRNA	linear	EST 18-JUN-2003
CD656661	LOCUS	CD656661	804 bp	mRNA	linear	EST 18-JUN-2003
	DEFINITION	AGENCOURT	14555875	NIA Human H1 Embryonic Stem Cell	CDNA Library	

ACCESSION	REFERENCE
VERSION	AUTHORS
KEYWORDS	TITLE
SOURCE	JOURNAL
ORGANISM	COMMENT

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

(Long) Homo sapiens cDNA clone IMAGE:30427288 5', mRNA sequence.
CD56661
CD56661.1 GI:31897180
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 804)
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabbs-r@mail.nih.gov
Tissue procurement: Irene Ginis and Mahendra Rao, NIA
cDNA library preparation: Yulan Piao and Minoru KO
cDNA library arrayed by: The I.M.A.G.E. Consortium (LNLML)
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC c lone distribution information
can be found through the I.M.A.G.E. Consortium/LNLML at:
<http://image.llnl.gov>
Plate: NDAM514 row: h column: 17
High quality sequence stop: 680.

FEATURES
source

1. accession="U04411" /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:30427288" /tissue_type="Embryonic Stem cells" /cell_line="WA01" /lab_host="DHI08 (T1 phage-resistant)" /clone_lib="NIA Human H1 Embryonic Stem Cell cDNA Library (Long)" /note=vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI; This is a long-transcript enriched cDNA library (Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]) from WA01 cell line. Undifferentiated human ES cell line WA01/H1 was obtained from WiCell Research Institute, Inc., Madison, WI, cultured according to their instructions, on MEF feeders. They formed round colonies with defined edges and were positive for alkaline phosphatase, SSEA-4, OCT3, OCT4, REX1, UTF, TERT, SOX2, CX34 and CX45. They are negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLT3, SSEA-1, TUBB3, NES, GFAP, and EOMES. When confluent (18-10 days after plating), the ES cells from 4 x 6 cm dishes were treated with 1 mg/ml collagenase, type IV (Invitrogen/GIBCO) for 5-10 min and gently scraped off with 5 ml pipette. RNA was purified with TRIzol Reagent from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558 (2001). [PMID:11544199] Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pGACTGTTCTAGATCGCAGCGCCGCTTTTTTTTTTTT-3'] from 3.4 µg of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lone-linker IL-Sal4, purified by phenol/chloroform extraction, and separated from free linkers by Centricon-100 column. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-s for 25 cycles. The products were purified by phenol/chloroform extraction and Centricon-100 column. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-Sport6 plasmid vector. The average insert size is about 3.6kb."

ORIGIN

Query Match	64.8%	Score 682.8;	DB 6;	Length 804;
Best Local Similarity	97.3%	Pred. No. 1.5e-170;		
Matches 716: Conservative	0;	Mismatches 17;	Indels 3;	Gaps 2


```
QY 212 TATCAAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAAGTACTTTCATCTCTC 271
DB 1 TATCAAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAAGTACTTTCATCTCTC 60
QY 272 ACATCTTAGC-CITTTTTCAGCCAGTGTGAATTTGAATGAATAATTTGGTGAGCGC 330
DB 61 ACATCTTAGCCTTTTTCAGCCAGTGTGAATTTGAATGAATAATTTGGTGAGCGC 120
QY 331 TTTAGTCAGAGGTGAGGTTCAGAGGCTCGCTGTTCTTATGGCTTTCAAATTTCTCATC 390
DB 121 TTTAGTCAGAGGTGAGGTTCAGAGGCTCGCTGTTCTTATGGCTTTCAAATTTCTCATC 180
QY 391 GAGATGTTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAGATCCCAAG 450
DB 181 GAGATGTTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAGATCCCAAG 240
QY 451 AAAAGGGAATTTTATTAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 510
DB 241 AAAAGGGAATTTTATTAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 300
QY 511 TGGGCTTCGGATGATAGATAGAAATCTACTTTTGGGGAAGAGTGGTGGCTTT 570
DB 301 TGGGCTTCGGATGATAGATAGAAATCTACTTTTGGGGAAGAGTGGTGGCTTT 360
QY 571 GCTGCTGTAGAAGGAGTTTCTTCTCAGGATCTTTTGTGCTATATTTCTGGCTAAAGAAG 630
DB 361 GCTGCTGTAGAAGGAGTTTCTTCTCAGGATCTTTTGTGCTATATTTCTGGCTAAAGAAG 420
QY 631 AGAGTCTTATGCCAGGATCACTTTTCCAAATGAATCAATCAGCAGAGATGAAGGACTT 690
DB 421 AGAGTCTTATGCCAGGATCACTTTTCCAAATGAATCAATCAGCAGAGATGAAGGACTT 480
QY 691 CACTGTGACTTGTGCTGCTGATGTTTCCAAATCTTGTAGTAAATGAAGCTTCAAGAAGAAG 750
DB 481 CACTGTGACTTGTGCTGCTGATGTTTCCAAATCTTGTAGTAAATGAAGCTTCAAGAAGAAG 540
QY 751 GTCAGGAGATCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 810
DB 541 GTCAGGAGATCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
QY 811 CCAGTTGGCTCATTTGAATGAATTTGATTTTGAATGAATTTGATGAATTTGATGAATTTGATG 870
DB 601 CCAGTTGGCTCATTTGATGAATTTGATTTTGAATGAATTTGATGAATTTGATGAATTTGATG 660
QY 871 GACAGATTTACTTGTGAATTTGATTTCTCAAGGTTTTCAGGAGAGAAATTCCTTTTGTAT 930
DB 661 GACAGATTTACTTGTGAATTTGATTTCTCAAGGTTTTCAGGAGAGAAATTCCTTTTGTAT 718
QY 931 TTTATGGAACAATTT 946
DB 719 TATGGAACAATTTT 734
```

RESULT 14

```
BM801298
LOCUS BM801298
DEFINITION AGENCOURT_6422860 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5586693
5', mRNA sequence.
ACCESSION BM801298
VERSION BM801298.1 GI:19118121
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1061)
NIH-MGC http://mgs.nhl.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
```

cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12354 row: p column: 22
High quality sequence stop: 631.

FEATURES
source

1..1061
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5586693"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_92"
/notes="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 62.1%; Score 653.8; DB 4; Length 1061;
Best Local Similarity 97.4%; Pred. No. 8.9e+163;
Matches 717; Conservative 0; Mismatches 13; Indels 6; Gaps 5;

QY 1 ATGGGGACCCGGAAAGCGCGGAGCGCGGGCTGATCAGGATGAGAGATCATCTTCA 60
DB 85 ATGGGGACCCGGAAAGCGCGGAGCGCGGGCTGATCAGGATGAGAGATCATCTTCA 144
QY 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAAGAGCCACTCTTAAAGAAAGTCTCGC 120
DB 145 GACACCAACGAAAGTGAATTAAGTCAAAATGAAAGAGCCACTCTTAAAGAAAGTCTCGC 204
QY 121 CGGTTTGTTCATCTTCCAAATCCAGTACCTGATATTTGGAAAATGTATAAACAGGCACAG 180
DB 205 CGGTTTGTTCATCTTCCAAATCCAGTACCTGATATTTGGAAAATGTATAAACAGGCACAG 264
QY 181 GCTTCTCTTGGACAGCAGAGAGGTTGATTTAAAGAGATCTCCCTCACTGGAACAAG 240
DB 265 GCTTCTCTTGGACAGCAGAGAGGTTGATTTAAAGAGATCTCCCTCACTGGAACAAG 324
QY 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTCTAGCCCTTTTGGAGCCAGTAT 300
DB 325 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTCTAGCCCTTTTGGAGCCAGTAT 384
QY 301 GGAATTTGAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCGAGGTTCCAGAGGCT 360
DB 385 GGAATTTGAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCGAGGTTCCAGAGGCT 444
QY 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCAGAAATGTTTCACTCAGAGATGTACAGTTTG 420
DB 445 CGCTGTTTCTATGGCTTTCAAATTTCTCATCAGAAATGTTTCACTCAGAGATGTACAGTTTG 504
QY 421 CTGATGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATGAA 480
DB 505 CTGATGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATGAA 564
QY 481 ACCATGCCCTATCTTAAGAAAAGCAGATTTGGCCCTTGGATGGATGAGATAGAAA 540
DB 565 ACCATGCCCTATCTTAAGAAAAGCAGATTTGGCCCTTGGATGGATGAGATAGAAA 624
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTAGAAAGGAGTTTCTTCTCAGGA 600
DB 625 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTAGAAAGGAGTTTCTTCTCAGGA 684
QY 601 TCTTTTGTGCTATATTTCTGGCTAAAGAAAGAGAGGTTTATGCGCAGGACTCACTTTTTC 660
DB 685 TCTTTTGTGCTATATTTCTGGCTAAAGAAAGAGAGGTC-TATGCCAGGACTCACTTTTTC 743
QY 661 AATGAACATCAT-CAGCAGAGAT-GAAGGACTTCACTGTGA--CTTTGCTTGCCTGAT-GT 715

```

Db      744 CATGGACTCATCCACGAGATGGAGGACTTCCCTGTGAACCTTTCCTTGCCTGTATGCT 803
Qy      716 TCCAATACCTAGTAAA 731
Db      804 TCCAATACCTAAGAAA 819

RESULT 15
BUI70979
LOCUS   BUI70979
DEFINITION BUI70979 896 bp mRNA linear EST 04-SEP-2002
5', mRNA sequence.
AGENCOURT_7940338 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6143754
ACCESSION BUI70979
VERSION   BUI70979.1 GI:22684963
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 896)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE    National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL  Unpublished (1999)
COMMENT  Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Cloned by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLMW13467 row: k column: 19
High quality sequence stop: 619.

FEATURES
source
1..896
Location/Qualifiers
1..896
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6143754"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_67"
/note="Organ: eye; Vector: pCMV-SPORT6; Site: 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 1.75 kb. Library constructed by Life
Technologies."

ORIGIN
Query Match 62.0%; Score 653.2; DB 5; Length 896;
Best Local Similarity 98.5%; Pred. No. 1.2e-162;
Matches 670; Conservative 0; Mismatches 8; Indels 2; Gaps 1;

Qy      1 ATGGGGACCCCGAAGCGCGGAGCGGCGGCTGGATCAGATCAGAGATCATCTTCA 60
Db      55 ATGGGGACCCCGAAGCGCGGAGCGGCGGCTGGATCAGATCAGAGATCATCTTCA 114
Qy      61 GACACCAACGAAAGTGAATAAGTCAATGAAGGCCACTCTTAAGAAAGAGTTCTCGC 120
Db      115 GACACCAACGAAAGTGAATAAGTCAATGAAGGCCACTCTTAAGAAAGAGTTCTCGC 174
Qy      121 CGGTTTGTATCTTCCATCCAGTACCTGATATTTGAAAATGTATATAACAGGCACAG 180
Db      175 CGGTTTGTATCTTCCATCCAGTACCTGATATTTGAAAATGTATATAACAGGCACAG 234
Qy      181 GCTTCCTCTGACAGCAGAGAGGTGACTTATCAAGGATCTCCCTCACTGGAAACAAG 240
Db      235 GCTTCCTCTGACAGCAGAGAGGTGCGACTTATCAAGGATCTCCCTCACTGGAAACAAG 294
Qy      241 CTTAAGCAGATGAGAAGTACTTCTCTCATCTTAGCCCTTTTTCAGCCAGTGAT 300
Db      295 CTTAAGCAGATGAGAAGTACTTCTCTCATCTTAGCCCTTTTTCAGCCAGTGAT 354

```

```

Qy      301 GGAATTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAGAGGTCAGGTTCCAGAGGCT 360
Db      355 GGAATTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAGAGGTCAGGTTCCAGAGGCT 414
Qy      361 CGCTGTTTCTATGGCTTTTCAAAATTTCTCATCAGAAATGTTCTACTCAGAGATGTACAGTTTG 420
Db      415 CGCTGTTTCTATGGCTTTTCAAAATTTCTCATCAGAAATGTTCTACTCAGAGATGTACAGTTTG 474
Qy      421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA 480
Db      475 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA 534
Qy      481 ACCATGCCCTTATGTTAAGAAAAAAGCAGATTGGGCCCTTGGATGGATAGCAGATAGAAAA 540
Db      535 ACCATGCCCTTATGTTAAGAAAAAAGCAGATTGGGCCCTTGGATGGATAGCAGATAGAAAA 594
Qy      541 TCTACTTTTGGGAAAAGAGTGGTGGCCTTTGCTGCTGTAGAAAGAGTTTCTTCTCAGGA 600
Db      595 TCTACTTTTGGGAAAAGAGTGGTGGCCTTTGCTGCTGTAGAAAGAGTTTCTTCTCAGGA 654
Qy      601 TCTTTGCTGCTATATTTCTGGCTAAAGAGAGGTCCTTATGCCAGGACTCCTTTT-- 658
Db      655 TCTTTGCTGCTATATTTCTGGCTAAAGAGAGGTCCTTATGCCAGGACTCCTTTTTC 714
Qy      659 CCAATGAACCTCATCAGCAGA 678
Db      715 CAATGAACCTCATCAGCAGA 734

```

Search completed: October 30, 2005, 03:23:17
Job time : 2721.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2005, 23:30:35 ; Search time 146 Seconds
(without alignments)
11801.366 Million cell updates/sec

Title: US-10-698-228-12

Perfect score: 1053

Sequence: 1 atggggcagccgaaagcc.....tcaccttgatgagatttt 1053

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/prodata/1/ina/5A COMB.seq.*
- 2: /cgn2_6/prodata/1/ina/5B COMB.seq.*
- 3: /cgn2_6/prodata/1/ina/6A COMB.seq.*
- 4: /cgn2_6/prodata/1/ina/6B COMB.seq.*
- 5: /cgn2_6/prodata/1/ina/6C COMB.seq.*
- 6: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1053	100.0	1053	4	US-10-019-733-12
2	1053	100.0	1081	4	US-10-019-733-4
3	1051.4	99.8	1053	4	US-10-019-733-2
4	1051.4	99.8	4955	4	US-10-019-733-3
5	580.6	55.1	2500	4	US-09-962-665-9
6	580.6	55.1	2500	4	US-09-963-333-9
7	580.6	55.1	2500	4	US-09-962-677-9
8	580	55.1	2500	4	US-09-023-655-1370
9	580	55.1	2500	4	US-09-949-016-145
10	578.4	54.9	2479	4	US-09-949-016-2025
11	439.4	41.7	481	3	US-08-905-223-125
12	429.8	40.8	14176	1	US-08-307-499-1
13	429.8	40.8	14176	1	US-08-307-499-14
14	429.8	40.8	14176	3	US-09-299-268-1
15	429.8	40.8	14176	3	US-09-299-268-14
16	384.8	36.5	1236	4	US-09-248-796A-3914
17	382.8	36.4	825	1	US-08-307-499-40
18	382.8	36.4	825	3	US-09-299-268-40
19	339.4	32.2	1112	1	US-08-136-743B-1
20	317.6	30.2	1083	4	US-09-248-796A-3915
21	316.6	30.1	419	3	US-08-905-223-129
22	260.8	24.8	656	3	US-09-328-111-791
23	195.2	18.5	449	4	US-09-270-767-14670
24	169.2	16.1	366	3	US-08-307-499-13
25	169.2	16.1	366	3	US-09-299-268-13
26	141.4	13.4	373	4	US-09-401-064-337
27	139.8	13.3	373	4	US-09-401-064-319

c	28	115.2	10.9	366	4	US-09-401-064-307	Sequence 307, App
	29	112.4	10.7	297	4	US-09-313-294A-4435	Sequence 4435, Ap
	30	93.8	8.9	11820	4	US-09-949-016-11887	Sequence 11887, A
	31	93.8	8.9	11826	4	US-09-949-016-13767	Sequence 13767, A
	32	92	8.7	601	4	US-09-949-016-20615	Sequence 20615, A
	33	92	8.7	601	4	US-09-949-016-69028	Sequence 69028, A
	34	76	7.2	276	4	US-09-313-294A-3437	Sequence 3437, Ap
	35	71.8	6.8	304	4	US-09-313-294A-7143	Sequence 7143, Ap
	36	68	6.5	276	4	US-09-313-294A-4638	Sequence 4638, Ap
	37	61	5.8	1053	4	US-09-902-540-2666	Sequence 2666, Ap
c	38	61	5.8	13706	4	US-09-902-540-1124	Sequence 1124, Ap
	39	60.8	5.7	1141	4	US-09-806-708B-22	Sequence 22, Appl
	40	60	5.7	601	4	US-09-949-016-69031	Sequence 69031, A
	41	54.4	5.2	288	4	US-09-313-294A-6108	Sequence 6108, Ap
	42	53.8	5.1	276	4	US-09-313-294A-3244	Sequence 3244, Ap
	43	50	4.7	1230025	4	US-09-198-452A-1	Sequence 1, Appli
	44	50	4.7	1230230	4	US-09-438-185A-1	Sequence 1, Appli
c	45	47	4.5	124884	4	US-09-661-596A-76	Sequence 76, Appl

ALIGNMENTS

RESULT 1

US-10-019-733-12
; Sequence 12, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619W00P
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 12
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-12

Query Match	100.0%;	Score 1053;	DB 4;	Length 1053;
Best Local Similarity	100.0%;	Pred. No. 6.8e-309;	Mismatches 0;	Indels 0;
Matches 1053;	Conservative 0;			Gaps 0;
Qy	1	ATGGGGACCCGGAAGCGCGGGCTGGATCAGGATGAGATCATCTTCA	60	
Db	1	ATGGGGACCCGGAAGCGCGGGCTGGATCAGGATGAGATCATCTTCA	60	
Qy	61	GACACCAACGAAAGTGAATTAAGTCAATGAAGCCACTCTCTAAGAAAGAGTTCTCGC	120	
Db	61	GACACCAACGAAAGTGAATTAAGTCAATGAAGCCACTCTCTAAGAAAGAGTTCTCGC	120	
Qy	121	CGTTTGTCTCTTCCATCCAGTACCTGATATTGGAAATGTATAACAGGCACAG	180	
Db	121	CGTTTGTCTCTTCCATCCAGTACCTGATATTGGAAATGTATAACAGGCACAG	180	
Qy	181	GCTTCTCTTGGACAGCAGAGAGTTGACTTATCAAGAGTCTCCCTCACTGGAACAAG	240	
Db	181	GCTTCTCTTGGACAGCAGAGAGTTGACTTATCAAGAGTCTCCCTCACTGGAACAAG	240	
Qy	241	CTTAAAGCAGATGAGAGTACTTCTCATCTCTACATCTTTTTCAGCCAGGTGAT	300	
Db	241	CTTAAAGCAGATGAGAGTACTTCTCATCTCTACATCTTTTTCAGCCAGGTGAT	300	
Qy	301	GGAAATGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGAGGTTCCAGAGGCT	360	

```
Db 301 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
QY 361 CGCTGTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTAAATGCAATTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTAAATGCAATTGAA 480
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTCGATGAGATAGCAGATAGAAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTCGATGAGATAGCAGATAGAAAA 540
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTCGCTGTGTAAGAGGATTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTCGCTGTGTAAGAGGATTTTCTTCTCAGGA 600
QY 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGCTTTATGCGAGGACTCACTTTTTC 660
Db 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGCTTTATGCGAGGACTCACTTTTTC 660
QY 661 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTTCTGCTGCTGATGTTCAA 720
Db 661 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTTCTGCTGCTGATGTTCAA 720
QY 721 TACTTAGTAATAAGCTTTCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAAA 780
Db 721 TACTTAGTAATAAGCTTTCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAAA 780
QY 781 ATTGAGCAGGAGTTTAAACAGAGGCTTGCAGATTGGCTCATTCAGGAATGAATGCAAT 840
Db 781 ATTGAGCAGGAGTTTAAACAGAGGCTTGCAGATTGGCTCATTCAGGAATGAATGCAAT 840
QY 841 TTGATGAAACAGTACATGAGTTTGTAGCTGACAGATTAATCTTGTGAACTTGGATTCTCA 900
Db 841 TTGATGAAACAGTACATGAGTTTGTAGCTGACAGATTAATCTTGTGAACTTGGATTCTCA 900
QY 901 AAGGTTTTTTCAGGAGAAAATCCTTTGATTTTATGAGAAAACATTTCTTTAGAGAAAA 960
Db 901 AAGGTTTTTTCAGGAGAAAATCCTTTGATTTTATGAGAAAACATTTCTTTAGAGAAAA 960
QY 961 ACAAAATTTCTTTGAGAAACGATTTTCAGAGTATCAGCGTTTTCAGAGTATGCGAGAAACC 1020
Db 961 ACAAAATTTCTTTGAGAAACGATTTTCAGAGTATCAGCGTTTTCAGAGTATGCGAGAAACC 1020
```

RESULT 2

```
US-10-019-733-4
; Sequence 4, Application US/10019733
; Patent No. 682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/019, 733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 4
; LENGTH: 1081
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-4
```

```
Query Match 100.0%; Score 1053; DB 4; Length 1081;
Best Local Similarity 100.0%; Pred. No. 6,9e-309;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ATGGGGCGCCCGGAAAGCGCGGCTGGATCAGATGAGAGATCATCTTCA 60
Db 20 ATGGGGCGCCCGGAAAGCGCGGCTGGATCAGATGAGAGATCATCTTCA 79
QY 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
Db 80 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 139
QY 121 CGGTTTGTCAATCTTCCAAATCCAGTACCTGATATTTGGAAATGTATAAACAGGCAAG 180
Db 140 CGGTTTGTCAATCTTCCAAATCCAGTACCTGATATTTGGAAATGTATAAACAGGCAAG 199
QY 181 GCTTCCTTCTGGACAGCAGAGAGGTTGACTTATCAAGGATCTCCCTCACTGGAACAAG 240
Db 200 GCTTCCTTCTGGACAGCAGAGAGGTTGACTTATCAAGGATCTCCCTCACTGGAACAAG 259
QY 241 CTTAAAGCAGATGAGAGTACTTCACTCTCACATCTTAGCCTTTTTCAGCCAGTAT 300
Db 260 CTTAAAGCAGATGAGAGTACTTCACTCTCACATCTTAGCCTTTTTCAGCCAGTAT 319
QY 301 GGAATTTGTAATAAGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAAGGCT 360
Db 320 GGAATTTGTAATAAGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAAGGCT 379
QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 420
Db 380 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 439
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTAAATGCAATTGAA 480
Db 440 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTAAATGCAATTGAA 499
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTCGATGAGATAGAGATAAGAAA 540
Db 500 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTCGATGAGATAGAGATAAGAAA 559
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTCGCTGTGTAAGAGGATTTTCTTCTCAGGA 600
Db 560 TCTACTTTTGGGGAAGAGTGGTGGCTTTCGCTGTGTAAGAGGATTTTCTTCTCAGGA 619
QY 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGCTTTATGCGAGGACTCACATTTTTC 660
Db 620 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGCTTTATGCGAGGACTCACATTTTTC 679
QY 661 AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTCTGCTGATGTTCCAA 720
Db 680 AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTCTGCTGATGTTCCAA 739
QY 721 TACTTAGTAATAAGCCTTCAGAAAGAGGCTCAGGAGATCATTTGTTGATGTTCAAA 780
Db 740 TACTTAGTAATAAGCCTTCAGAAAGAGGCTCAGGAGATCATTTGTTGATGTTCAAA 799
QY 781 ATTGAGCAGGAGTTTAAACAGAGGCTTGCAGATTGGCTCATTTGGAATGAATGCAAT 840
Db 800 ATTGAGCAGGAGTTTAAACAGAGGCTTGCAGATTGGCTCATTTGGAATGAATGCAAT 859
QY 841 TTGATGAAACAGTACATGAGTTTGTAGCTGACAGATTAATCTTGTGAACTTGGATTCTCA 900
Db 860 TTGATGAAACAGTACATGAGTTTGTAGCTGACAGATTAATCTTGTGAACTTGGATTCTCA 919
QY 901 AAGGTTTTTTCAGGAGAAAATCCTTTGATTTTATGAGAAAACATTTCTTTAGAGAAAA 960
Db 920 AAGGTTTTTTCAGGAGAAAATCCTTTGATTTTATGAGAAAACATTTCTTTAGAGAAAA 979
QY 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGAGTATGCGAGAAACC 1020
```

Db 980 ACAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGATTATGGCAGAAACC 1039
QY 1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053
Db 1040 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1072

RESULT 3
US-10-019-733-2
; Sequence 2, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-2

Query Match 99.8%; Score 1051.4; DB 4; Length 1053;
Best Local Similarity 99.9%; Pred. No. 2.1e-308;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGGACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 ATGGCGGACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
QY 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
QY 121 CGGTTTGTCTTTCCTCAATCCAGTACCGTGTATTTGGAAATGTATAAACAGGCACAG 180
Db 121 CGGTTTGTCTTTCCTCAATCCAGTACCGTGTATTTGGAAATGTATAAACAGGCACAG 180
QY 181 GCTTCCTTCTGGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGACACAG 240
Db 181 GCTTCCTTCTGGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGACACAG 240
QY 241 CTTAAAGCAGATGAGAAGTACTTCTCTCATCTCTAGCGCTTTTTCAGCCAGTGTAT 300
Db 241 CTTAAAGCAGATGAGAAGTACTTCTCTCATCTCTAGCGCTTTTTCAGCCAGTGTAT 300
QY 301 GGAATTGTAATGAAATTTGGTGAGCGCTTTTACTCAGGAGGTGAGGTTCCAGAGGCT 360
Db 301 GGAATTGTAATGAAATTTGGTGAGCGCTTTTACTCAGGAGGTGAGGTTCCAGAGGCT 360
QY 361 CGCTGTTCTATGGCTTTCAATTTCTCATCGAGATGTTCTCATGAGATGATACAGTTTG 420
Db 361 CGCTGTTCTATGGCTTTCAATTTCTCATCGAGATGTTCTCATGAGATGATACAGTTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTAAATCAATTTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTAAATCAATTTGAA 480
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTTTCGATGGATGAGCAGATAGAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTTTCGATGGATGAGCAGATAGAAA 540

QY 541 TCTACTTTTGGGAAAGAGTGGTGGCCCTTTGCTGCTGTAGAGAGAGTTTCTTCTCAGA 600
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCCCTTTGCTGCTGTAGAGAGAGTTTCTTCTCAGA 600
QY 601 TCTTTTGTCTGTATATCTGGCTAAAGAGAGAGGTTCTTATGCCAGGACTCCTTTTTC 660
Db 601 TCTTTTGTCTGTATATCTGGCTAAAGAGAGAGGTTCTTATGCCAGGACTCCTTTTTC 660
QY 661 AATGAACTCATCAGCAGAGATGAAGGACTTCTCCTGTGACTTTTGTGCTGTATGTTCCAA 720
Db 661 AATGAACTCATCAGCAGAGATGAAGGACTTCTCCTGTGACTTTTGTGCTGTATGTTCCAA 720
QY 721 TACTTAGTAAATAAGCCTTCAGAAAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 780
Db 721 TACTTAGTAAATAAGCCTTCAGAAAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 780
QY 781 ATTGAGCAGGAGTTTAAACAGAAAGCCTTCCAGAGTTGGCCCTCATTTGGAATGGAATT 840
Db 781 ATTGAGCAGGAGTTTAAACAGAAAGCCTTCCAGAGTTGGCCCTCATTTGGAATGGAATT 840
QY 841 TTGATGAAACAGTACATTTAGCTGTAGCTGACAGATTTCTTTGGAACCTTTGGATTCTCA 900
Db 841 TTGATGAAACAGTACATTTAGCTGTAGCTGACAGATTTCTTTGGAACCTTTGGATTCTCA 900
QY 901 AAGGTTTTTCAGGACAGAAATCCTTTTGAATTTTATGAAACATTTCTTTAGAGGAAAA 960
Db 901 AAGGTTTTTCAGGACAGAAATCCTTTTGAATTTTATGAAACATTTCTTTAGAGGAAAA 960
QY 961 ACAATTTCTTTCAGAAACAGGTTTTCAGAGTATCAGCGTTTTCAGGATGAGCAGAAACC 1020
Db 961 ACAATTTCTTTCAGAAACAGGTTTTCAGAGTATCAGCGTTTTCAGGATGAGCAGAAACC 1020
QY 1021 ACAGATAACGCTTCTTCACTTGGATGAGATTTT 1053
Db 1021 ACAGATAACGCTTCTTCACTTGGATGAGATTTT 1053

RESULT 4
US-10-019-733-3
; Sequence 3, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 3
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-3

Query Match 99.8%; Score 1051.4; DB 4; Length 4955;
Best Local Similarity 99.9%; Pred. No. 5.1e-308;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGGACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 245 ATGGCGGACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
QY 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAAGAAAGAGTTCTCGC 120
Db 305 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAAGAAAGAGTTCTCGC 364

```
QY 121 CGTTTGTTCATCTTCCCAATCCAGTACCTGTATATTGGAAATGTATAAACAGGCACAG 180
Db |||||
QY 365 CGTTTGTTCATCTTCCCAATCCAGTACCTGTATATTGGAAATGTATAAACAGGCACAG 424
Db |||||
QY 181 GCTTCCTTCTGACAGCAGAGAGAGGTGTGATCTATCAAGGATCTCCCTCACTGGAAACAG 240
Db |||||
QY 425 GCTTCCTTCTGACAGCAGAGAGAGGTGTGATCTATCAAGGATCTCCCTCACTGGAAACAG 484
QY 241 CTTAAGCAGATGAGAGTACTTCACTCTCACTCTTAGCCCTTTTTCGACCCAGTGT 300
Db |||||
QY 485 CTTAAGCAGATGAGAGTACTTCACTCTCACTCTTAGCCCTTTTTCGACCCAGTGT 544
QY 301 GGAATTTGTAATGAAAAATTTGGTGAGCGCTTTTAGTCAGAGGTGCAGGTTCACAGGCT 360
Db |||||
QY 545 GGAATTTGTAATGAAAAATTTGGTGAGCGCTTTTAGTCAGAGGTGCAGGTTCACAGGCT 604
QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAATGTTTCACTCAGAGATGACAGTTTG 420
Db |||||
QY 605 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAATGTTTCACTCAGAGATGACAGTTTG 664
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAAATTTTATTTAATGCAATTGAA 480
Db |||||
QY 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAAATTTTATTTAATGCAATTGAA 724
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTCGCGATGGATGACAGATAGAAAA 540
Db |||||
QY 725 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTCGCGATGGATGACAGATAGAAAA 784
QY 541 TCTACTTTTGGGAAAGAGTGGTGCCCTTGTGCTGTAGAGAGGATTTTCTTCACAGA 600
Db |||||
QY 785 TCTACTTTTGGGAAAGAGTGGTGCCCTTGTGCTGTAGAGAGGATTTTCTTCACAGA 844
QY 601 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTACTTTTTC 660
Db |||||
QY 845 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTACTTTTTC 904
QY 661 AATGAACATCAGCAGAGATCAAGGACTTCACTGTGACTTTGCTGCTGATGTTCCAA 720
Db |||||
QY 905 AATGAACATCAGCAGAGATCAAGGACTTCACTGTGACTTTGCTGCTGATGTTCCAA 964
QY 721 TACTTAGTAAATTAAGCCCTTCAGAGAAAGGCTCAGGAGATCATTTGATGCTGTCAA 780
Db |||||
QY 965 TACTTAGTAAATTAAGCCCTTCAGAGAAAGGCTCAGGAGATCATTTGATGCTGTCAA 1024
QY 781 ATTGAGCAGAGTGTAAACAGAGCCTTGCAGTTGGCTCATTTGGAATGAATTTGCATT 840
Db |||||
QY 1025 ATTGAGCAGAGTGTAAACAGAGCCTTGCAGTTGGCTCATTTGGAATGAATTTGCATT 1084
QY 841 TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATCTCA 900
Db |||||
QY 1085 TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATCTCA 1144
QY 901 AAGGTTTTTTCAGGCAGAAAAATCCTTTGATTTTATGGAACAATTTCTTAGAAGGAAA 960
Db |||||
QY 1145 AAGGTTTTTTCAGGCAGAAAAATCCTTTGATTTTATGGAACAATTTCTTAGAAGGAAA 1204
QY 961 ACAAAATTTCTTTGAGAACAGGATTTTTCAGAGTATCAGCGTTTTCGAGTTATGCGAGAAC 1020
Db |||||
QY 1205 ACAAAATTTCTTTGAGAACAGGATTTTTCAGAGTATCAGCGTTTTCGAGTTATGCGAGAAC 1264
QY 1021 ACAGATAAGCTTTTTCACCTTGGATGAGATTTT 1053
Db |||||
QY 1265 ACAGATAAGCTTTTTCACCTTGGATGAGATTTT 1297
```

RESULT 5

```
US-09-962-665-9
; Sequence 9, Application US/09962665
; Patent No. 6537759
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: POLYLPOLYGLUTAMATE SYNTHETASE GENE SEQUENCE
```

```
; TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING THE
; FILE REFERENCE: 11926-015004
; CURRENT APPLICATION NUMBER: US/09/962,665
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 128, 1464
; OTHER INFORMATION: n = g or a
; NAME/KEY: misc_feature
; LOCATION: 189
; OTHER INFORMATION: n = t or g
; NAME/KEY: misc_feature
; LOCATION: 524
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc_feature
; LOCATION: 1399
; OTHER INFORMATION: n = t or a
; NAME/KEY: misc_feature
; LOCATION: 1636, 1738, 2259
; OTHER INFORMATION: n = c or t
; US-09-962-665-9
```

```
Query Match 55.1%; Score 580.6; DB 4; Length 2500;
Best Local Similarity 75.1%; Pred. No. 2.2e-165;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

QY 89 ATGAGAGCCACTCTAAGAAAGAGTTCTCGCCGGTTTGTCACTTTTCCAAATCCAGTACC 148
Db |||||
QY 397 AGATAGCGCGTCTGTGAGAGAAACCCCGCGCTTTGTCACTTTCCCAATCCAGTACC 456
QY 149 CTGATATTTGGAAAAATGTATAACAGGCACAGGCTTCTCTTCTGGACAGAGAGGTTG 208
Db |||||
QY 457 ATGATATCTGGCAGATGTATAGAAAGCAGAGGCTTCTTTTGGACCGCGAGAGGTTG 516
QY 209 ACTTATCAAGAGATCTCCCTCACTGGAAACAAGCTTTAAAGCAGATGAGAACTTCACTCT 268
Db |||||
QY 517 ACCTCTCNAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGATATTTTATAT 576
QY 269 CTCACATCTTAGCCTTTTTCGAGCCAGTGTGAATTTGAAATGAAATTTTGGTGAGC 328
Db |||||
QY 577 CCATGTTCTGGCTTTCTTTTGCAGCAAGCGATGGCATAGTAAATGAAACCTTGGTGAGC 636
QY 329 GCTTTAGTCAGAGAGTGCAGGTTCCAGAGGCTCGCTGTTTCTATGGCTTTTCAAAATTTCTCA 388
Db |||||
QY 637 GATTTAGCCAGAGATTCAGATTACAGAGCCCGCTGTTTCTATGGTTCGAAATTTGCCA 696
QY 389 TCGAGAAATGTTCACTCAGAGATGTACAGTTTTCGTGATAGACACTTACATCAGAGATCCCA 448
Db |||||
QY 697 TGGAAAAACATACATTCTGAAATGTATAGTCTTCTTATTGACACTTACATAAAGATCCCA 756
QY 449 AGAAAGGGAATTTTATTTAATGCAATTTGAAACCATGCCCTATGTTTAAAGAAAAAGCAG 508
Db |||||
QY 757 AAGAAAGGGAATTTTCTTCAATGCCAATTTGAAACGATGCCCTTGTGTCAAGAAAGCAG 816
QY 509 ATTGGGCTTGGATGGATGATAGAGATAGAAAAATCTACTTTTGGGAAAGAGTGTGGCT 568
```

```
Db 817 ACTGGGCTTGGCTGGATGGGGACAAAGAGGCTACTATGGTGAACGTGTGTAGCCT 876
QY 569 TTGCTGCTGTAGAAGAGTCTTCTCAGGATCTTTTGGCTGTATATCTGTGCTAAGA 628
Db 877 TTGCTGCTGTAGAAGCAATTTCTTTCGGTCTTTTGGCTGTATATCTGTGCTCAAGA 936
QY 629 AGAGAGGCTTATGGCAGGACTCATTCTTTCATGAAGTCAATCAGCAGAGATGAAGGAC 688
Db 937 AACGAGGACTGATGCTGCGCTCACATTTCTTAATGAATTTATAGCAGATGAGGTT 996
QY 689 TTCACCTGTGACTTTGCTTGGCTGTATGCTTCAATACCTTAGTAATAGCCTTCAAGAA 748
Db 997 TACACTGTGATTTGCTTGGCTGTATGCTTCAATACCTTAGTAATAGCCTTCAAGAA 1056
QY 749 GGTGAGGAGATCATTTGTTGATGCTTCAATACCTTAGTAATAGCCTTCAAGAA 808
Db 1057 GAGTAGAAGAAATATATCAATGCTTTCGGATAGAACAGGAGTTCCTCAGTGGCCT 1116
QY 809 TGCAGTGTGGCTCATTTGGAATGAATTTGATGAACAGTACATTTGATTTGTAG 868
Db 1117 TGCCTGTGAGCTCATTTGGAATGAATTTGATGAACAGTACATTTGATTTGTAG 1176
QY 869 CTGACAGATTAATTTGGAATTTGATGAACAGTACATTTGATTTGTAG 928
Db 1177 CAGACAGACTTATGCTGGAATTTGATGAACAGTACATTTGATTTGTAG 1236
QY 929 ATTTATGGAACATTTTCTTATGGAATTTGATGAACAGTACATTTGATTTGTAG 988
Db 1237 ACTTTATGGAATTTTCTTATGGAATTTGATGAACAGTACATTTGATTTGTAG 1296
QY 989 AGTATCAGGATTTGATGGAATTTGATGAACAGTACATTTGATTTGTAG 1048
Db 1297 AGTATCAGGATTTGATGGAATTTGATGAACAGTACATTTGATTTGTAG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360
```

RESULT 6

```
US-09-963-333-9
; Sequence 9, Application US/09963333
; Patent No. 6664062
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: THYMIDINE SYNTHASE GENE SEQUENCE VARIANCES
; TITLE OF INVENTION: HAVING UTILITY IN DETERMINING THE TREATMENT
; TITLE OF INVENTION: OF DISEASE
; FILE REFERENCE: 11926-015002
; CURRENT APPLICATION NUMBER: US/09/963,333
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 128, 1464
; OTHER INFORMATION: n = g or a
; NAME/KEY: misc_feature
; LOCATION: 189
```

```
; OTHER INFORMATION: n = t or g
; NAME/KEY: misc_feature
; LOCATION: 524
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc_feature
; LOCATION: 1399
; OTHER INFORMATION: n = t or a
; NAME/KEY: misc_feature
; LOCATION: 1636, 1738, 2259
; OTHER INFORMATION: n = c or t
US-09-963-333-9
```

```
Query Match 55.1%; Score 580.6; DB 4; Length 2500;
Best Local Similarity 75.1%; Pred. No. 2.2e-165;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;
```

```
QY 89 ATGAAGAGCCACTCTTAAGAAAGATTTCTGCGCGTTTGTCTATCTTTTCCAATCCAGTACC 148
Db 397 AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTTGTCTATCTTTCCCATCGAGTACC 456
QY 149 CTGATATTTGAAATATTAACAGCAGCAGGCTTCTTCTGACAGCAGAGAGGTTG 208
Db 457 ATGATATCTGSCAGATATAAGAAAGCAGAGGCTTCTTTTGGACCGCCGAGGAGTTG 516
QY 209 ACTTATCAAGAGTCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAAGTACTTCACT 268
Db 517 ACCTCTCNAAGGACATTCAGCACCTGGGAATCCCTGAAACCCGAGAGAGATATTTTATAT 576
QY 269 CTCACATCTTATGCGCTTTTTCAGCAGCAGTATGAATTTGTAATGAAATTTGGTGGAGC 328
Db 577 CCCATGTTCTGGCTTCTTTTGCAGCAAGCATGTCATAGTAATGAAATTTGGTGGAGC 636
QY 329 GCTTTAGTCAGGAGGTCAGGTTCCAGAGGCTGCTGTTCTATGCTGTTTCAATTTCTCA 388
Db 637 GATTAGCCAAGAAGTTTCAATTAACAAGCCCGCTTCTTATGGCTTCCAAATTTGCA 696
QY 389 TCGAGATGTTCACTCAGAGATGTACAGTGTGCTGATAGACACTTACATCAGAGATCCCA 448
Db 697 TGAAGAACATACATTTCTGAATGTATAGTCTTCTTATGACACTTACATAAAGATCCCA 756
QY 449 AGAAAGGGAATTTTATTTAATGAATTTGAACCATGCCCCCTATGTTTAAAGAAAAAGCAG 508
Db 757 AGAAGAGGGAATTTCTTCAATGCCATTTGAACGATGCTTGTGTCAAGAAAGAGGAGC 816
QY 509 ATTGGGCTTTCGATGATGATAGAGATAAATCTACTTTTGGGAAAGAGTGGGCT 568
Db 817 ACTGGGCTTTCGCTGATTTGGGACAAAGAGGCTACTATGTTGAACCTGTTGTAGCCT 876
QY 569 TTGCTGCTGTAGAAGGATTTTCTTCTCAGATCTTTTGTGCTATATTTCTGGCTAAGA 628
Db 877 TTGCTGAGTGGAAAGGCAATTTCTTTTCCGGTTCCTTTTGGCTCGATATTTCTGGCTCA 936
QY 629 AGAGAGGCTTATGTCAGGAGTCACTCTTTTCCAAATGAATCACTCAGCAGAGATGAAGGAC 688
Db 937 AACGAGGACTGATGCTGGCTCACATTTTCTAATGAATTTATGACAGATGAGGTT 996
QY 689 TTCACCTGTGATTTGCTTGGCTGATTTCAATCTATGTAATTAAGCTTTTCAAGAA 748
Db 997 TACACTGTGATTTTGTCTGCTGATTTCAACACACCTGGTACACAAACCATCGGAGGAGA 1056
QY 749 GGTGAGGAGATCATTTGTTGATGCTGTCAAAATTTGAGCAGGAGTTTTAAACAGAGCT 808
Db 1057 GAGTAAGAGAAATATATCAATGCTTTCGGATAGAACAGGAGTTCCTCAGTGGGCT 1116
QY 809 TGCAGTGTGGCTCATTTGGAATGAATTTGATGAACAGTACATTTGATTTGTAG 868
Db 1117 TGCCTGTGAGCTCATTTGGAATGAATTTGATGAACAGTACATTTGATTTGTAG 1176
QY 869 CTGACAGATTTACTTTGGAATTTGGAATTTTCAAGGTTTTCAGGCGAGAAATCTTTT 928
Db 1177 CAGACAGACTTATGCTGGAATTTGAGGAGTTCAGAGTTCAGAGTTCAGAGTTCAG 1236
QY 929 ATTTATGGAACATTTCTTTTGAAGAAAAACAATTTCTTTTGAAGAACAGGATTTT 988
```

Db 1237 ACTTTATGAGAATATTTCACTGGAAGGAAAGACTAACTTCTTTGAGAAGAGTAGGCG 1296
QY 989 AGTATCAGCGTTTGCAGTTATGGCAAGAACACACAGATAACGCTCTTCACTTGGATGCAG 1048
Db 1297 AGTATCAGAGGATGGGAGTGATGTCAGTCCACAGAGAATCTTTTACCTTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 7

US-09-962-677-9
; Sequence 9, Application US/09962677
; Patent No. 6759200
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: THYMIDINE PHOSPHORYLASE GENE SEQUENCE
; TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING
; TITLE OF INVENTION: THE TREATMENT OF DISEASE
; FILE REFERENCE: 11926-015003
; CURRENT APPLICATION NUMBER: US/09/962,677
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 128..1464
; OTHER INFORMATION: n = g or a
; NAME/KEY: misc_feature
; LOCATION: 189
; OTHER INFORMATION: n = t or g
; NAME/KEY: misc_feature
; LOCATION: 524
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc_feature
; LOCATION: 1399
; OTHER INFORMATION: n = t or a
; NAME/KEY: misc_feature
; LOCATION: 1636..1738, 2259
; OTHER INFORMATION: n = c or t
US-09-962-677-9

Query Match 55.1%; Score 580.6; DB 4; Length 2500;

Best Local Similarity 75.1%; Pred. No. 2.2e-165;

Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

QY 89 ATGAGAGCCATCTCTAAGAAAGAGTTCTGCGCGGTTTGTCATCTTTTCAATCCAGTACC 148
Db 397 AGGATGAGCGCTGTGAGAGAAACCCCGCGCTTTGTCTCTTTCCCATCGAGTACC 456
QY 149 CTGATATTTGGAATGTATTAACAGGCACAGCTTCTTCTGGACAGAGAGGTTG 208
Db 457 ATGATATCTGGCAGATGTATAGAGAGGAGAGGCTTCTTTTGGACCGCGGAGGTTG 516
QY 209 ACTTATCAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAAGTACTTCACT 268
Db 517 ACCTCTCAAGACATTCAGCACTGGGATCCTGAAACCGGAGAGATATTTATAT 576

QY 269 CTCACATCTTAGCCTTTTGTGAGCCAGTGAATTTGTAATAAATTTGTTGGAGC 328
Db 577 CCATGTTCTGGCTTTTCTTTGAGCAAGCGATGCAATAGTAATAAATTTGTTGGAGC 636
QY 329 GCTTTAGTCAGAGAGTGCAGGTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAAATTTCA 388
Db 637 GATTTAGCCAAGAAGTTTCAGATTTACAGAAGCCCGCTGTTTCTATGGCTTCCAAATTTGCA 696
QY 389 TCGAGAATGTTCACTCAGAGATGTACAGATTTGCTGTAGACACTTACATCAGAGATCCCA 448
Db 697 TGGAAAAACATACATTTCTGAAATGTATAGTCTTCTTATTCACACTTACATATAAAGATCCCA 756
QY 449 AGAAAAAGGGAATTTTATTTAATGCAATTTGAAACCATGCCCCCTATGTTAAGAAAAAAGCAG 508
Db 757 AGAAAAAGGGAATTTTCTTCAATGCCATTTGAAACGATGCCCCCTTGTGTCAAGAAGAGGAG 816
QY 509 ATTGGGCTTTCGATGATAGCAGATAGAAAAATCTACTTTTGGGGAAGAGTGGTGGCT 568
Db 817 ACTGGGCTTTCGCTGATTTGGGGAACAAGAGGCTACCTATGTTGAACGTTGTGTAGCCT 876
QY 569 TTGCTGCTGTAGAAGAGTTTCTTCTCAGGATCTTTTGTGCTGCTATATTTCTGGCTAAGA 628
Db 877 TTGCTGCTGAGTGAAGGCAATTTCTTTTCCGCTTCTTTTGGCTCGATATTTCTGGCTCAAG 936
QY 629 AGAGAGGCTCTTATGCCAGGACTCACATTTTCCAAATGAACCTCATCAGCAGAGATGAAGGAC 688
Db 937 AACGAGGACTGATGCTGGCCTCACATTTTCTAATGAACCTTATAGCAGAGATGAGGGTT 996
QY 689 TTCACCTGTGACTTTGCTTGGCTGATGTTCCAATCTTAGTAAATTAAGCCTTCAGAAGAAA 748
Db 997 TACACTGTGATTTGCTTGGCTGATGTTCAACACCTGGTTACACAAACCATCGAGGAGA 1056
QY 749 GGGTCAGGAGATCATTTGATGCTGTCAAAATTTGAGCAGGAGTGTTTTAAACAGAGCCT 808
Db 1057 GAGTAAGAGAAATAATTAATCAATGCTGTTCCGATAGAACACAGGAGTTCTCTCACTGAGGCT 1116
QY 809 TGCCAGTTGGCCTCATTTGGAATGAATTTGATTTGATGAACAGTACATTCAGTTTGTAG 868
Db 1117 TGCTGTGAAGCTCATTTGGGATGAATTTGCACTCTAATGAAGCAATACATTCAGTTTGTGG 1176
QY 869 CTGACAGATTTACTTTGTGGAACCTTTGGATTTCTCAAGGTTTTCAGGCAGAAAAATCCTTTG 928
Db 1177 CAGACAGACTTATGCTGGAACCTGGGTTTGTAGCAAGGTTTTCAGAGTAGAGACCAATTTG 1236
QY 929 ATTTTATGGAACAATTTCTTTAGAGGAAAAACAATTTCTTTGAGAAAAAGAGTTTTCAG 988
Db 1237 ACTTTATGGAGATATTTCACTGGAAGGAAAGACTAACTTCTTTGAGAAGAGATAGGCG 1296
QY 989 AGTATCAGCGTTTTCAGATTTATGGCAGAAACACAGATACGTTCTTCACTTGGATGCAG 1048
Db 1297 AGTATCAGAGGATGGGAGTGATGTCAGTCCAAACAGAGAAATTTCTTTTACCTTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 8

US-09-023-655-1370

; Sequence 1370, Application US/09023655

; Patent No. 6607879

; GENERAL INFORMATION:

; APPLICANT: Cocks, Benjamin G.

; APPLICANT: Susan G. Stuart

; APPLICANT: Jeffrey J. Seilhamer

; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE

; TITLE OF INVENTION: EXPRESSION

; NUMBER OF SEQUENCES: 1508

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

; STREET: 3174 PORTER DRIVE

; CITY: PALO ALTO

STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/023,655
FILING DATE: HERWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0001 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1370:
SEQUENCE CHARACTERISTICS:
LENGTH: 2500 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g36154
US-09-023-655-1370

Query Match 55.1%; Score 580; DB 4; Length 2500;
Best Local Similarity 75.1%; Pred. No. 3.4e-165;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

QY 89 ATGAAGAGCCATCTCTAAGAAAGAGTTCTGCGGGTTGTGCTATCTTCCAAATCCAGTACC 148
DB 397 AGGATGAGCGCTGCTGAGAGAAAACCCCGCGCTTGTCTATCTTCCCATCGAGTACC 456
QY 149 CTGATATTGGAAAATGTATAACAGCACAGGCTTCTTCTTGACAGAGAGAGGTTG 208
DB 457 ATGATATCGGAGATGTATAAGAGGAGGCTTCTTGTGACCGCCGAGGAGGTTG 516
QY 209 ACTTATCAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAAGTCTTCAATCT 268
DB 517 ACTCTCCAAGGACATTCAGCACTGGGAATCCTGAAACCCGAGGAGAGATATTTATAT 576
QY 269 CTCACATCTTAGCCTTTTGGAGCAGTGATGGAATTTGTAATGAAAATTTGGTGGAGC 328
DB 577 CCCATGTTCTGGCTTCTTGGCAGCAAGCGATGGCATAGTAAATGAAAACCTTGGTGGAGC 636
QY 329 GCTTTAGTCAGAGTGCAAGTTCCAGAGCTCGCTGTTTCTATGCTTCCAAATCTCA 388
DB 637 GATTTAGCCAAAGTTCAGATACAGAGCCGCTGTTTCTATGCTTCCAAATGCCA 696
QY 389 TCGAGATGTTCACTCAGAGATGTACAGTTTGTGATAGACACTTACATCAGAGATCCCA 448
DB 697 TGAANAACATATCTGAAATGTAGTCTTATTGACACTTACATAAAGATCCCA 756
QY 449 AGAAAGGGAATTTTATTATTAATGCAATTTGAAACCAATGCCCTATGTTAAGAAAAAGCAG 508
DB 757 AAGAAAGGGAATTTCTTCAATGCAATTTGAAACGATGCTTGTGTCAGAAAGAGGAGCAG 816
QY 509 ATTGGGCTTCGATGATAGCAGATAGAAATCTACTTTTGGGNAAGAGTGGTGGCT 568
DB 817 ACTGGGCTTCGCTGGATTTGGGGAACAAAGAGGCTTACCTATGTTGAACTGTTGAGCT 876
QY 569 TTGCTGCTGTAAGAGGAGTTTCTTCTCAGGATCTTTTCTGCTATATTCTGCTTAAAGA 628
DB 877 TTGCTGAGTGAAGGCATTTTCTTTCGGTCTTTTGGCTCGATATTCTGGCTCAAGA 936

QY 629 AGAGAGGTCTTATGCCAGGACTCACTTTTCCAAATGAACCTCATCAGCAGAGATGAAGAC 688
DB 937 AACGAGGACTGATGCTGGCTCACATTTTCTAATGAACCTTATTAGCAGAGATGAGGTT 996
QY 689 TTCACGTGTGATTTGCTTGGCTGATGTTCCAAATCTTAGTAAATAGCCTTTCAGAAAGAA 748
DB 997 TACACTGTGATTTGCTTGGCTGATGTTCAAAACACCTGGTACACAAACCATCGGAGGAGA 1056
QY 749 GGCCTCAGGAGATCATTTGATGCTCTCAAAATTTGACGAGGTTTTTAACAGAGCCT 808
DB 1057 GAGTAAGAGAAATAAATTATCAATGCTGTTCGATAGAACAGGAGTTCTCACTGAGGCT 1116
QY 809 TGCAGTGTGGCTCATTTGGAATGAATTCATTTTATGAAACAGTACATTCAGTTTGTAG 868
DB 1117 TGCTGTGAAGCTCATTTGGATGATTCACCTTAATGAACAATACATTCAGTTTGTG 1176
QY 869 CTGACAGATTACTTTGTGGAATCTTGGATTTCTCAAGGTTTTTTCAGGCAGAAATCTTTTG 928
DB 1177 CAGACAGACTTATGCTGGAATCTGTTTTCAGCAAGGTTTTTTCAGAGTAGAGAACCCATTG 1236
QY 929 ATTTTATGGAACAATTTCTTTAGAGGAAAAACAAATTTCTTTGAGAAACGAGTTTCAG 988
DB 1237 ACTTTATGGAGAAATTTTCTGGAAGGAAAGACTAACTTCTTTGAGAAAGAGAGTAGGCG 1296
QY 989 AGTATCAGCGTTTTCAGATTTATGCGAGAAACACAGATAACCTTTCACCTTGGATGCG 1048
DB 1297 AGTATCAGAGATGGGAGTGTCAAGTCCAAAGAGAAATCTTTTACCTTGAATGCTG 1356
QY 1049 ATTT 1052
DB 1357 ACTT 1360

RESULT 9

US-09-949-016-145
; Sequence 145, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 145
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-145

Query Match 55.1%; Score 580; DB 4; Length 2500;
Best Local Similarity 75.1%; Pred. No. 3.4e-165;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

QY 89 ATGAAGAGCCATCTCTAAGAAAGAGTTCTGCGGGTTGTGCTATCTTCCAAATCCAGTACC 148
DB 397 AGGATGAGCGCTGCTGAGAGAAAACCCCGCGCTTGTCTATCTTCCCATCGAGTACC 456
QY 149 CTGATATTGGAAAATGTATAACAGCACAGGCTTCTTCTTGACAGAGAGAGGTTG 208
DB 457 ATGATATCGGAGATGTATAAGAGGAGGCTTCTTGTGACCGCCGAGGAGGTTG 516
QY 209 ACTTATCAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAAGTCTTCACT 268

Db 517 ACCTCTCCAGGACATTCACGACTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT 576
QY 269 CTCACATCTTAGCTTTTTCGAGCCAGTGATGGAATTCGTAATGAAATTTGGTGAGC 328
Db 577 CCATGCTCTGGCTTTCTTTCGACGAGGATGCGCATAGTAATGAAACTTTGGTGAGC 636
QY 329 GCTTTAGTCAGGAGTGCGAGTTTCCAGAGGCTCGCTGTTTCTATGCTTTTCAAAATCTCA 388
Db 637 GATTTAGCCAAGAGTTTCAGATTACAGAACCGCGCTGTTTCTATGCTTCCAAATGGCA 696
QY 389 TCAGAAATGTTCACTCAGAGATGTAAGTTTGTGATAGACACTTACATCAGAGATCCCA 448
Db 697 TCGAAACATACATCTCGAAATGATAGTCTTCTTATGACACTTACATCAGAGATCCCA 756
QY 449 AGAAAGGGAATTTTATTAATGCAATTTGAACCAATGCAATGCAATGCAATGCAATGCA 508
Db 757 AAGAAAGGGAATTTTCTTCAATGCAATTTGAACCAATGCAATGCAATGCAATGCAATGCA 816
QY 509 ATTGGGCTTTGGGATGATAGCAGATAGAAATCTACTTTTGGGGAAGAGTGGTGGCCT 568
Db 817 ACTGGGCTTTGGGATGATAGCAGATAGAAATCTACTTTTGGGGAAGAGTGGTGGCCT 876
QY 569 TTGCTGCTGAGAGAGTTTCTTCTCAGGATCTTTTCTGCTGCTATATTTCTGGCTAAAGA 628
Db 877 TTGCTGCTGAGAGAGTTTCTTCTTCCGGTCTTTTCTGCTGCTATATTTCTGGCTCAAGA 936
QY 629 AGAGAGTCTTATGCGAGGACTCACTTTTCCAAATGCAATGCAATGCAATGCAATGCAATGCA 688
Db 937 AACGAGGATGATGCGCTGCGCTCACATTTTCTTAATGAACCTTATAGCAGAGATGAGGGT 996
QY 689 TTCACTGTGACTTTGCTTGGCTGATGCTTCCAAATCTACTTTTGAATAGGCTTTCAGAAAGAA 748
Db 997 TACACTGTGATTTTGGCTTGGCTGATGCTTCCAAATCTACTTTTGAATAGGCTTTCAGAAAGAA 1056
QY 749 GGGTCAGGAGATCAITTTGATGCTTCAAAATGAGCAGGAGTGTTCCTTAAACAGAGCCT 808
Db 1057 GAGTAGAGAAATTAATTAATCAATGCTGTTGCGGATAGACAGGAGTCTCTACTGAGGCT 1116
QY 809 TGCAGTGTGGCTCATTTGGAAATGAAATGCAATTTTGTATGAAACAGTACATGAGTTGTAG 868
Db 1117 TGCCTGTGAAGCTCATTTGGGATGAAATGCAATTTGCACTCTAATGAAGCAATACATGAGTTGTGG 1176
QY 869 CTGACAGATTTACTTTGGAACTTTGGATTTCTCAAGGTTTTCAGGCGAGAAATCCTTTTG 928
Db 1177 CAGACAGACTTATGCTGGAACTGGGTTTATAGCAAGGTTTTCAGAGTGAAGAACCCATTTG 1236
QY 929 ATTTTATGAAACATTTCTTTTGAAGGAAACAAATTTCTTTGAGAAACGAGTTTCAG 988

RESULT 10
US-09-949-016-2025
; Sequence 2025, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2025
; LENGTH: 2479
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-2025

Query Match 54.9%; Score 578.4; DB 4; Length 2479;
Best Local Similarity 75.0%; Pred. No. 1e-164;
Matches 723; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

QY 89 ATGAAGAGCCACTCTTAAGAAAGATTCGCGCGTTTGTTCATCTTCCATCCAGTACC 148
Db 397 AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTTGTTCATCTTCCCATCCAGTACC 456
QY 149 CTGATATTTGGAAATGTATAAACAGCAGCAGGCTTCTTCTCGACAGCAGAGAGTTG 208
Db 457 ATGATATCTGGCAGATGTATAAGAGCAGAGGCTTCTTGTGACCGCCGAGAGGTTG 516
QY 209 ACTTATCAAGAGATCTCCCTCACTGGAAACAAGCTTAAAGCAGATGAGAAGTACTTCACT 268
Db 517 ACCTCTCCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT 576
QY 269 CTCACATCTTAGCTTTTTCGAGCCAGATGATGAAATGTAATGAAATTTGTTGAGC 328
Db 577 CCATGTTCTGGCTTTCTTTTGCAGCAAGCGATGGCATAGTAAATGAAATCTTGTGGAGC 636
QY 329 GCTTTAGTCAGGAGTGAGGTTCCAGAGGCTCGCTGTTTCTATGCTTTCATGCTTCAATCTCA 388
Db 637 GATTTAGCCAAGAGTTTCAGATTACAGAACCGCTGTTTCTATGGCTTCCAAATGGCA 696
QY 389 TCAGAAATGTTCACTCAGAGATGTACAGATTTTGTGTATGACACTTACATCAGAGATCCCA 448
Db 697 TCGAAACATACATCTGAAATGATAGTCTTCTTATGACACTTACATCAGAGATCCCA 756
QY 449 AGAAAGGGAATTTTATTAATGCAATTTGAACCAATGCAATGCAATGCAATGCAATGCA 508
Db 757 AGAAAGGGAATTTTCTTCAATGCAATTTGAACCAATGCAATGCAATGCAATGCAATGCA 816
QY 509 ATTGGGCTTTGGGATGATAGCAGATAGAAATCTACTTTTGGGGAAGAGTGGTGGCCT 568
Db 817 ACTGGGCTTTGGGATGATAGCAGATAGAAATCTACTTTTGGGGAAGAGTGGTGGCCT 876
QY 569 TTGCTGCTGAGAGAGTTTCTTCTCAGGATCTTTTCTGCTGCTATATTTCTGGCTAAAGA 628
Db 877 TTGCTGCTGAGAGAGTTTCTTCTTCCGGTCTTTTCTGCTGCTATATTTCTGGCTCAAGA 936
QY 629 AGAGAGTCTTATGCGAGGACTCACTTTTTCCAATGCAATGCAATGCAATGCAATGCAATGCA 688
Db 937 AACGAGGATGATGCGCTGCGCTCACATTTTCTTAATGAACCTTATAGCAGAGATGAGGGT 996
QY 689 TTTCACTGTGACTTTGCTTGGCTGATGCTTCCAAATCTACTTTTGAATAGGCTTTCAGAAAGAA 748
Db 997 TACACTGTGATTTTGGCTTGGCTGATGCTTCCAAATCTACTTTTGAATAGGCTTTCAGAAAGAA 1056
QY 749 GGGTCAGGAGATCAITTTGATGCTTCAAAATGAGCAGGAGTGTTCCTTAAACAGAGCCT 808
Db 1057 GAGTAGAGAAATTAATTAATCAATGCTGTTGCGGATAGACAGGAGTCTCTACTGAGGCT 1116
QY 809 TGCAGTGTGGCTCATTTGGAAATGAAATGCAATTTTGTATGAAACAGTACATGAGTTGTAG 868
Db 1117 TGCCTGTGAAGCTCATTTGGGATGAAATGCAATTTGCACTCTAATGAAGCAATACATGAGTTGTGG 1176
QY 869 CTGACAGATTTACTTTGGAACTTTGGATTTCTCAAGGTTTTCAGGCGAGAAATCCTTTTG 928
Db 1177 CAGACAGACTTATGCTGGAACTGGGTTTATAGCAAGGTTTTCAGAGTGAAGAACCCATTTG 1236
QY 929 ATTTTATGAAACATTTCTTTTGAAGGAAACAAATTTCTTTGAGAAACGAGTTTCAG 988

Db 1237 ACTTTATGAGAATATTTCTACTGGAAGGAAAGACTAACTTTCTTTGAGAAGAGTAGGCG 1296
QY 989 AGTATCAGCGTTTTCAGTTATGGCAGAAACACACAGATAACGCTTTCACCTTGGATGCAG 1048
Db 1297 AGTATCAGGAGTGGAGTGAATGCTAAGTCCAAACAGAGAATCTTTTACCTTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 11

US-08-905-223-125
; Sequence 125, Application US/08905223
; Patent No. 622029
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste D.
; APPLICANT: Duleurt, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS
; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobb, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Win95
SOFTWARE: Word

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/905,223
FILING DATE:
CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 125:

SEQUENCE CHARACTERISTICS:
LENGTH: 481 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: DOUBLE
TOPOLOGY: LINEAR
MOLECULE TYPE: CDNA
ORGANISM: Homo Sapiens
TISSUE TYPE: Brain
FEATURE:

NAME/KEY: sig_peptide
LOCATION: 41..343
IDENTIFICATION METHOD: Von Heijne matrix
OTHER INFORMATION: score 4.4
OTHER INFORMATION: seq ISHILAFPAASDG/IV

Query Match 41.7%; Score 439.4; DB 3; Length 481;
Best Local Similarity 99.8%; Pred. No. 5.9e-123;
Matches 440; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAGCGCGGAGCGGCGGCTGGATCAGATGAGATCATCTTCA 60
Db 41 ATGGGCGACCCGGAAGCGCGGAGCGGCGGCTGGATCAGATGAGATCATCTTCA 100
QY 61 GACACCAAGGAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAGAGTTCTCGC 120
Db 101 GACACCAAGGAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAGAGTTCTCGC 160

QY 121 CGTTTGTCTATCTTTTCAATCCAGTACCTGATATTTGGAAAAATGTATAAACAGGCACAG 180
Db 161 CGTTTGTCTATCTTTTCAATCCAGTACCTGATATTTGGAAAAATGTATAAACAGGCACAG 220
QY 181 GCTTCTCTTGGACAGCAGAGAGGTTGACTTATCAAAAGGATCTCCCTCATCTGGAAACAAG 240
Db 221 GCTTCTCTTGGACAGCAGAGAGGTTGACTTATCAAAAGGATCTCCCTCATCTGGAAACAAG 280
QY 241 CTTAAGCAGATGAGAAGTACTTCTCTCATCTTCTAGCCTTTTTCAGCCAGGTAT 300
Db 281 CTTAAGCAGATGAGAAGTACTTCTCTCATCTTCTAGCCTTTTTCAGCCAGGTAT 340
QY 301 GGAATTGTAATGAAAAATTTGGTGGAGCGCTTTTGTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
Db 341 GGAATTGTAATGAAAAATTTGGTGGAGCGCTTTTGTAGTCAGGAGGTGCAGGTTCCAGAGGCT 400
QY 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db 401 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 460
QY 421 CTGATAGACACTTACATCAGA 441
Db 461 CTGATAGACACTTACATCAGA 481

RESULT 12

US-08-307-499-1/c
; Sequence 1, Application US/08307499
; Patent No. 5651972
; GENERAL INFORMATION:

APPLICANT: Moyer, Richard W.
APPLICANT: Vi uela, Eladio

TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
TITLE OF INVENTION: Live Vaccine Vector

NUMBER OF SEQUENCES: 60

CORRESPONDENCE ADDRESS:

ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: U.S.A.

ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/307,499

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/908,241

FILING DATE: 1-JUL-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/908,630

FILING DATE: 29-JUN-1992

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/342,212

FILING DATE: 21-APR-1992

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Saliwanchik, David R.

REGISTRATION NUMBER: 31,794

REFERENCE/DOCKET NUMBER: UF35.1.FWCCI

TELECOMMUNICATION INFORMATION:

TELEPHONE: 904-375-8100

TELEFAX: 904-372-5800

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

```
;
; LENGTH: 14176 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3852..4226
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 4585..4887
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5131..5310
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5760..5912
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 6786..7130
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 10148..10513
;
US-08-307-499-1
```

```
Query Match 40.8%; Score 429.8; DB 1; Length 14176;
Best Local Similarity 66.2%; Pred. No. 3.3e-119;
Matches 551; Conservative 0; Mismatches 327; Indels 5; Gaps 2;

QY 71 AAAGTGAATAAAGTCAAAATGAAGACCCACTCCTAAGAAAGAGTTCTCCCGGTTGTCA 130
Db 10841 AAAATGAATTTTATACACGAATGGAGCCCTATTCTTCAAGAGCTGATTTCTAGGTTGCTTA 10782

QY 131 TCTTCCATCCAGTACCTGATATTGGAATATGTAAGACGACAGGCTTCTTCT 190
Db 10781 TTTTCCCTATTAAAGTATCATGATATCTGGAATATGTAAGACGAGTGGCAAGTTT 10722

QY 191 GGACAGCAGACGAGGTGACTTATCAAGAGATCTCCCTCACTGGCAACAGCTTAAAGCAG 250
Db 10721 GGACGTTGAAGAGTAGATTTATCAAGAGATTAGATGATGGGATAAATTAACCTAAG 10662

QY 251 ATGAGAGTACTTCACTCTCACTTTAGCCCTTTTTCAGCCAGTGTAGGAATGTAA 310
Db 10661 ACGAAAATACCTTTATAAAACATATACCTAGCATTTTTCATCTAGTGATGTATTGTAA 10602

QY 311 ATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCTCGCTGTTCT 370
Db 10601 ATGAGAATTTAGCGGAAAGATTTATGTGGATGTACAGTGTTCAGAGGCGACGATGTTCT 10542

QY 371 ATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTGTGTGATAGACA 430
Db 10541 ATGGATTTCAAAATGCTATGGAATATTTCACTCAGAAATGATAGTTATTATTAATAGATA 10482

QY 431 CTATCATCAGAGATCCCAAGAAAGGGAATTTTATTTAATCAATGAAACCATGCCCT 490
Db 10481 CATATGTAAGAGATAATATAGAAAATATGCAATTTATTTAAGCTATAGAAAATGGAAT 10422

QY 491 ATGTTAAGAAAACAGAGATTGGGCTTCGAGTGTAGTACAGATAGAAAATCTACTTTTG 550
Db 10421 GCGTAAAAAAGAAAGCTGATTTGGGCGCAGAAAATGGAATATC---TAGCAACAAGGATATG 10365

QY 551 GCGAAGAGTGTGGCTTTGCTGCTGTGTAGAGGAGTTTCTTCTCAGGATCTTTTGCTG 610
Db 10364 GAGAAGAGTAGTAGCAITTTGACGCTGTGGAGGGAATATTTCTTTCTGTTTCAITTTGCTG 10305

QY 611 CTATATCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTCCAAATGAACTCA 670
Db 10304 CTATATTTGGATAAAAAACAGGATTCATGCCGGAATACATTTTCTAATGAACTAA 10245

QY 671 TCAGCAGATGAAGGACTTCACTGTGACTTTGCTTGTGCTGATGTTTCCAAATCTACTAGTAA 730
Db 10244 TAAGTAGACACGAAGGTTTACATTTGATTTGGCGTGTAAATGTTTAAACATTTATTATAC 10185
```

RESULT 13

```
US-08-307-499-14
; Sequence 14, Application US/08307499
; Patent No. 5651972
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; TITLE OF INVENTION: Live Vaccine Vector
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: U.S.A.
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/307,499
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,241
; FILING DATE: 1-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,630
; FILING DATE: 29-JUN-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/342,212
; FILING DATE: 21-APR-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF35.1.FWCC1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14176 base pairs
```

TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:

NAME/KEY: CDS
LOCATION: 138..1460
FEATURE:
NAME/KEY: CDS
LOCATION: 2456..2659
FEATURE:
NAME/KEY: CDS
LOCATION: 2809..3030
FEATURE:
NAME/KEY: CDS
LOCATION: 3070..3330
FEATURE:
NAME/KEY: CDS
LOCATION: 3356..4180
FEATURE:
NAME/KEY: CDS
LOCATION: 4392..5894
FEATURE:
NAME/KEY: CDS
LOCATION: 6171..6398
FEATURE:
NAME/KEY: CDS
LOCATION: 6447..6875
FEATURE:
NAME/KEY: CDS
LOCATION: 6928..7431
FEATURE:
NAME/KEY: CDS
LOCATION: 7454..7858
FEATURE:
NAME/KEY: CDS
LOCATION: 7895..8155
FEATURE:
NAME/KEY: CDS
LOCATION: 8215..8682
FEATURE:
NAME/KEY: CDS
LOCATION: 8715..9539
FEATURE:
NAME/KEY: CDS
LOCATION: 9562..10272
FEATURE:
NAME/KEY: CDS
LOCATION: 10316..11908
FEATURE:
NAME/KEY: CDS
LOCATION: 11971..12780
FEATURE:
NAME/KEY: CDS
LOCATION: 12829..13107
FEATURE:
NAME/KEY: CDS
LOCATION: 13149..14171
US-08-307-499-14

Query Match 40.8%; Score 429.8; DB 1; Length 14176;
Best Local Similarity 66.2%; Pred. No. 3.3e-119;
Matches 651; Conservative 0; Mismatches 327; Indels 5; Gaps 2;

QY 71 AAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGCCGGTTTGTC 130
Db 3336 AAAATGATTTTATACAGAAATGGAGCCCTATTCTTCAAGAGTCTGATTTAGTTCGTTA 3395

QY 131 TCTTTTCAATCCAGTACCTCGATATTTGGAAATGTATAAAGCAGGCAAGGCTTCTTCT 190
Db 3396 TTTTCCCTATTAGTATCATGATATCTGGAAATGTATAAAGCAATCAGTGGCAAGTTTTT 3455

QY 191 GGACAGCAGAGAGGTTGACTTATCAAGAGTCTCCCTCACTGGAAACAAGCTTAAGCAG 250

Db 3456 GGACCGTTGAAGAGTAGATTATCAAAAGATTTAGATGATTGGGATAAATTAACCTAAAG 3515
QY 251 ATGAGAAAGTACTTTCATCTCTCACATCTTAGCTTTTGGAGCCAGTGTATGGAATGTAA 310
Db 3516 ACAGAAAATATCTTTTATAAAACATATCTAGCATTTTGGCATCTAGTGTATGTATAA 3575
QY 311 ATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGTTCCAGAGGCTGCTGTTTCT 370
Db 3576 ATGAGAAATTTAGCGAAAGATTTTATGTGGATGTACAGTGTTCAGAGGCAAGATTTCT 3635
QY 371 ATGGCTTTCAAAATCTCATCGAGAATGTTCATCTCAGAGATGTACAGTTTGTCTGATAGACA 430
Db 3636 ATGGAITTTCAATAGCTATGGAATATTCATTCAGAAATGTATAGTTTATTAATAGATA 3695
QY 431 CTTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTTGAAACCATGCCCT 490
Db 3696 CATATGTAGAGATAATATAGAAAAAATGCAATTTTATTTAAGCTATAGAAAACAATGGAAT 3755
QY 491 ATGTTAAGAAAAAGCAGATTTGGCCCTTGGCATGATAGATAGAGATAAGAAATCTACTTTG 550
Db 3756 GCGTAAAAAAGAAAGCTGATTTGGGCCAGAAATGGATATC---TAGCAACAAGGTATATG 3812
QY 551 GGGAAAGAGTGGTGGCCCTTTGCTCTGTAGAAGAGATTTTCTCTCAGGATCTTTTGTGCTG 610
Db 3813 GAGAAAGAGTAGTAGCATTTTGCAGCTGTGGGGAATATCTTTTCTGTTCAATTTGCTG 3872
QY 611 CTATATTTGCGCTAAAGAGAGAGGTCTTATGSCAGAGCTCACTTTTCCAAATGAACCTCA 670
Db 3873 CTATATTTTGGATAAAAAAAGCAGAGATTGATGCCGGATTAAACATTTTCTAATGAACATA 3932
QY 671 TCAGCAGAGATGAAGGACTTCACTGTGCACTTTGCTTGCCTGATGTTCCAATACTTTAGTAA 730
Db 3933 TAAAGTAGAGACGAAGGTTTACATTTGCTGTTTTCGCTGTGTTAATGTTTAAACAATTTATTAC 3992
QY 731 ATAAGCCTTCAGAAAGAGGTTCAGGAGAGATCATTTGTTGATGCTGCTCAAAATTTGAGCAGG 790
Db 3993 ATCCACCATCTAAGAAAGTTATACGTCGATATCATTTGATGCGGTTAATATAGAAAAGG 4052
QY 791 AGTTTTTAAACAGAGCCTTGCAGTTCGCCCTCATTTGGAATGAATTCGCAATTTTGTGTAAC 850
Db 4053 AGTTTTTGACAGTTGCTTATTCGGTGGATCTTATAGGTATGAATTTGTTGTTAATGTCTC 4112
QY 851 AGTACATTTAGTTTGTAGCTGACAGATTTCTTGTGGAACTTGTGGATTTCTCAAGGTTTTTC 910
Db 4113 AGTATATAGAATTCGTCCAGATAGATTATTAACAGAGTTAGGTTGTGAAAAG--TCTCA 4170
QY 911 AGGACAGAAATCCTTTTGTATTTATGAAAACATTTCTTTAGAGGAAACAAATTTCT 970
Db 4171 ATGTATATATCTTTTATGCTTTATGAGATATATATCACTAGAAGGTAAAGTAAATTTT 4230
QY 971 TTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTATTCGAGTATTCGCAAGAACACAGATAACG 1030
Db 4231 TCGAACGAGGTTAGTGAATATCAAAAGATGGGGTGTGTTTACAAATTAAGAAAGATA 4290
QY 1031 TCTTCACTTTGGATGCGATTTT 1053
Db 4291 TATTTTCTACGGATATAGATTTT 4313

RESULT 14
US-09-299-268-1/c
; Sequence 1, Application US/09299268
; Patent No. 6217882
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; TITLE OF INVENTION: Live Vaccine Vector
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik

STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: U.S.A.
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/299,268
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/901,127
FILING DATE:
APPLICATION NUMBER: US 07/908,241
FILING DATE: 1-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/908,630
FILING DATE: 29-JUN-1992
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/342,212
FILING DATE: 21-APR-1992
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Salliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UF35.1.FWCC1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 14176 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 3852..4226
FEATURE:
NAME/KEY: CDS
LOCATION: 4585..4887
FEATURE:
NAME/KEY: CDS
LOCATION: 5131..5310
FEATURE:
NAME/KEY: CDS
LOCATION: 5760..5912
FEATURE:
NAME/KEY: CDS
LOCATION: 6786..7130
FEATURE:
NAME/KEY: CDS
LOCATION: 10148..10513
US-09-299-268-1

Query Match 40.8%; Score 429.8; DB 3; Length 14176;
Best Local Similarity 66.2%; Pred. No. 3.3e-119;
Matches 651; Conservative 0; Mismatches 327; Indels 5; Gaps 2;

QY	71	AAAGTAAATAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTCTCGCGCGGTTGTCA	130
DB	10841	AAATGATTTTATACACGAATGGAGCCTTCTTCAAGAGTCTGATCTAGGTTGTTA	10782
QY	131	TCTTTTCAATCCAGTACCTCGATATTTGGAAAATGATATAACAGGCACAGGCTTCCTTCT	190
DB	10781	TTTTCCCTATTAAATCATGATATCTGGAAAATGATATAACCAATCAGTGGCAAGTTTTT	10722

QY	191	GGACAGCAGAAAGAGTTGACTTATCAAAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAG	250
DB	10721	GGACCGTTGAAGAAGTAGATTTTATCAAAAGATTTAGATGATTGGGATAAATTAACATAAG	10662
QY	251	ATGAGAAGTACTTCATCTCTCACATCTTAGCCTTTTGGAGCCAGTGTAGGAATGTAA	310
DB	10661	ACGAAAAATACTTTATAAAAACATATACCTAGCATTTTGGCATCTAGTGTATGTAA	10602
QY	311	ATGAAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAAGTTTCCAGAGGCTCGCTGTTCT	370
DB	10601	ATGAGAATTTAGCGGAAAGATTTTATGTGATGTACAGTGTTCAGAGCAGATGTTTCT	10542
QY	371	ATGCGTTTCAAATTTCTCATCGAATGTTTCACTCAGAGATGTACAGTTTGTGTATAGACA	430
DB	10541	ATGGAATTTCAAATAGCTATCGAAAAATATTCATTAGAAAAATGATAGTTTATTAATAGATA	10482
QY	431	CTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTTGAAACCATGCGCT	490
DB	10481	CATATGTAGAGATAATATAGAAAAAATGCAATTTTATTTAAGCGTATAGAAAAAATGGAAT	10422
QY	491	ATGTTAAGAAAAAGCAGATTTGGGCGCTTGGCATGTAGAGATAGAGAAAAATCTACTTTTG	550
DB	10421	GCCTAAAAAAGAAAGCTGATTTGGGCCAGAAAAATGGATATC---TAGCAACAGGTATATG	10365
QY	551	GGGAAAGAGTGTGGCGCTTTTGTCTGTAGAAAGAGTCTTCTCTCAGAGATCTTTTGTCTG	610
DB	10364	GAGAAAGAGTAGTAGCATTTTGCAGCTGTGGGGAATATCTTTTCTGTTCAITTTGCTG	10305
QY	611	CTATATTTCTGGCTAAAGAGAGAGGTCTTATGCGCAGGACTCACATTTTCCAAATGAACCTCA	670
DB	10304	CTATATTTTGGATAAAAAAAGCAGGATTTGATGCGCGGATTAACATTTTCTAATGAACATA	10245
QY	671	TCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCAATACTTAGTAA	730
DB	10244	TAAGTAGAGACGAAGGTTTACATTTGCGTGTAAATGTTTAAACATTTTATTTATAC	10185
QY	731	ATAAGCCTTCAGAAAGAAAGGTCAGGAGAGTCAATTTGTTGATGCTGTCAAAATTCAGCAGG	790
DB	10184	ATCCACCATCTAAGGAAGTTATTAACGTCGATATCATTTGATCGGTTAATATAGAAAAGG	10125
QY	791	AGTTTTTAACAGAAAGCCTTCCAGTGTGGCGCTCATTTGGAATGAATTTGATTTGATGAAC	850
DB	10124	AGTTTTTGACAGTTGCTTATTCGCGTGGATCTTATAGGTATGAATTTGTTTAAATGTCTC	10065
QY	851	AGTACATTTGAGTTTGTAGCTGCAGAGATTAATTTGGAACTTGGATTTCTCAAGGTTTTTC	910
DB	10064	AGTATATAGAAATTCGTCGAGATAGATTTATTAACAGAGTTAGGTTGTGAAAAG--TCTCA	10007
QY	911	AGGCAAAAAATCCTTTTGTATTTTATGAAAAACATTTCTTTAGAAAGAAAAACAAATTTCT	970
DB	10006	ATGTATATATCTTTTGTAGCTTTTATGGAGTATATATCACTAGAGGTAAAGACTAATTTT	9947
QY	971	TTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTTTATGGCAGAAACCAACAGATAACG	1030
DB	9946	TCGAACGACGAGTTAGTGAATATCAAAAGATGGGGTGTTTTACAAAATAAGAGAGAGATA	9887
QY	1031	TCTTCACTTGGATGCAGATTTT	1053
DB	9886	TATTTTCTACGGATATAGATTTT	9864

RESULT 15
US-09-299-268-14
; Sequence 14, Application US/09299268
; Patent No. 6217882
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; TITLE OF INVENTION: Live Vaccine Vector
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:


```
QY 731 ATAAGCCTTCAGAAGAAAGGTCAGGAGATCATTTGATGCTGTCAAAATTTAGCAGG 790
Db |||||
3993 ATCCACATCTAAGGAAGTTATAACGTCGATAATCATTCATGCGGTTAATATAGAAAGG 4052
QY 791 AGTTTTTAACAGAAGCCTTGCCAGTTGGCCTCATTTGGAATGAATTCATTTTGATGAAAC 850
Db |||||
4053 AGTTTTTGAAGTTGCTATTCGGGTGGATCTTATAGGTATGAATTTGTTTAAATGCTC 4112
QY 851 AGTACATTGAGTTTGTAGCTGACAGATTACTTTGTGGAACTTGGATTCTCAAAGGTTTTTC 910
Db |||||
4113 AGTATATAGANTCGTCGAGATAGATTATTAAACAGAGTTAGTTGTGAAAG--TCTCA 4170
QY 911 AGGCAGAAATCCTTTTGTATTTTATGGAATAACATTTCTTTAGAGGAAAAACAAATTTCT 970
Db |||||
4171 ATGTATATAATCCTTTTAGCTTTATCGAGTATATATCACTAGAGGTAAGACTAATTTT 4230
QY 971 TTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACCAACAGATAACG 1030
Db |||||
4231 TCGAACGACGAGTTAGTGAATATCAAAAGATGGGGGTGTTTACAAATAAAGAGAGATA 4290
QY 1031 TCTTCACCTTGGATGCAGATTTT 1053
Db |||||
4291 TATTTTCTACGGATATAGATTTT 4313
```

Search completed: October 30, 2005, 03:28:15
Job time : 149 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 30, 2005, 01:51:37 ; Search time 676.5 Seconds
(without alignments)

12855.457 Million cell updates/sec

Title: US-10-698-228-12

Perfect score: 1053

Sequence: 1 atggcgaccgcgaaggcc.....tcacctggatgcagatttt 1053

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 9784742 seqs, 4129495052 residues

Total number of hits satisfying chosen parameters: 19569484

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
3:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
4:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
5:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
6:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
7:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
8:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
9:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
10:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
11:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
12:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
13:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
14:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
15:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
16:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
17:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
18:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
19:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
20:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
21:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
22:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
23:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
24:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
25:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
26:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
27:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
28:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1053	100.0	1053	19	US-10-698-228-12	Sequence 12, Appl
2	1053	100.0	1053	19	US-10-698-228-4	Sequence 4, Appl
3	1051.4	99.8	1051	19	US-10-698-228-2	Sequence 2, Appl
4	1051.4	99.8	4955	18	US-10-172-118-71	Sequence 71, Appl
5	1051.4	99.8	4955	19	US-10-342-887-71	Sequence 71, Appl

6	1051.4	99.8	4955	19	US-10-698-228-3	Sequence 3, Appl
7	580	55.1	2482	24	US-10-220-335-514	Sequence 724, App
8	580	55.1	2500	9	US-09-954-456-724	Sequence 1169, Ap
9	580	55.1	2500	9	US-09-954-456-1169	Sequence 1827, Ap
10	580	55.1	2500	9	US-09-954-456-1827	Sequence 1370, Ap
11	580	55.1	2500	19	US-10-641-643-1370	Sequence 458, App
12	580	55.1	2500	21	US-10-733-878-458	Sequence 3751, Ap
13	580	55.1	2500	22	US-10-843-641A-458	Sequence 4196, Ap
14	580	55.1	2500	22	US-10-843-641A-4854	Sequence 4854, Ap
15	580	55.1	2500	22	US-10-756-149-713	Sequence 713, App
16	580	55.1	2500	24	US-09-925-301-505	Sequence 505, App
17	578.4	54.9	1989	9	US-10-084-817-342	Sequence 342, App
18	578.4	54.9	2216	16	US-10-084-817-342	Sequence 1636, Ap
19	570.4	54.2	2113	22	US-10-764-420-1636	Sequence 170, App
20	569	54.0	2641	24	US-10-220-335-170	Sequence 75, Appl
21	562.6	53.4	1328	19	US-10-403-571-75	Sequence 790, App
22	545.4	51.8	977	18	US-10-264-237-790	Sequence 15278, A
23	514	48.8	1371	24	US-10-450-763-15278	Sequence 19181, A
24	463	44.0	1289	26	US-11-097-143-19181	Sequence 34, Appl
25	447.8	42.5	186854	22	US-10-872-156-34	Sequence 174, Appl
26	442	42.0	1146	9	US-09-822-830A-174	Sequence 7245, Ap
27	433.8	41.2	1218	16	US-10-128-714-7245	Sequence 19180, A
28	420.4	39.9	3945	26	US-11-097-143-19180	Sequence 34052, A
29	404.8	38.4	963	18	US-10-369-493-34052	Sequence 2245, Ap
30	398.6	37.9	1146	16	US-10-128-714-2245	Sequence 46252, A
31	394.2	37.4	1173	18	US-10-369-493-46252	Sequence 6537, Ap
32	393.6	37.4	1242	17	US-10-032-585-6537	Sequence 126843,
33	388.2	36.9	1292	18	US-10-369-493-27809	Sequence 40471, A
34	382	36.3	1200	19	US-10-424-599-126843	Sequence 73024, A
35	380.6	36.1	1248	19	US-10-424-599-40471	Sequence 43, Appl
36	372.8	35.4	1361	19	US-10-424-599-73024	Sequence 1245, Ap
37	365.8	34.7	1586	21	US-10-853-047-43	Sequence 6245, Ap
38	363.2	34.5	957	18	US-10-320-797-2145	Sequence 245, App
39	363.2	34.5	1206	16	US-10-128-714-1245	Sequence 5245, Ap
40	350.6	33.3	1314	16	US-10-128-714-1245	Sequence 6851, Ap
41	350.6	33.3	1450	16	US-10-128-714-245	Sequence 12381, A
42	350.6	33.3	3314	16	US-10-128-714-245	
43	350.6	33.3	3450	16	US-10-128-714-5245	
44	349.8	33.2	629	24	US-10-450-763-6851	
45	328.2	31.2	560	9	US-09-864-761-12381	

ALIGNMENTS

RESULT 1

US-10-698-228-12
; Sequence 12, Application US/10698228
; Publication No. US20040072253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619W00P
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 12
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-12

Query Match

100.0%; Score 1053; DB 19; Length 1053;

Best Local Similarity 100.0%; Pred. No. 9.3e-295;		Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1	ATGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATCAGATCATCTTCA	60
Db	1	ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATCAGATCATCTTCA	60
Qy	61	GACACCAACGAAAGTGAATAAAGTCAATGAAGACCACTCTTAAGAAAGAGTCTCGC	120
Db	61	GACACCAACGAAAGTGAATAAAGTCAATGAAGACCACTCTTAAGAAAGAGTCTCGC	120
Qy	121	CGGTTTGTCACTTTCCTCAATCCAGTACCTGTATTTTGGAAATGATATAACAGGACAG	180
Db	121	CGGTTTGTCACTTTCCTCAATCCAGTACCTGTATTTTGGAAATGATATAACAGGACAG	180
Qy	181	GCTTCTCTCTGACACAGAGAGGTGACTTATCAAGGATCTCCCTCACTGGAACAG	240
Db	181	GCTTCTCTCTGACACAGAGAGGTGACTTATCAAGGATCTCCCTCACTGGAACAG	240
Qy	241	CTTAAAGCAGATGAGAAGTACTTCTCTCACTCTTACATCTTACATCTTACATCTTAC	300
Db	241	CTTAAAGCAGATGAGAAGTACTTCTCTCACTCTTACATCTTACATCTTACATCTTAC	300
Qy	301	GGAAATGTAATGAAATTTGGTGGAGCGCTTTTGTAGTACAGAGTTCAGAGGCT	360
Db	301	GGAAATGTAATGAAATTTGGTGGAGCGCTTTTGTAGTACAGAGTTCAGAGGCT	360
Qy	361	CGCTGTTCTATGGCTTTCCTCAATCTCATCGAGAGTGTTCCTCACTGAGATGATCAG	420
Db	361	CGCTGTTCTATGGCTTTCCTCAATCTCATCGAGAGTGTTCCTCACTGAGATGATCAG	420
Qy	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATGAA	480
Db	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATGAA	480
Qy	481	ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCCGATGATAGAGATAGAAA	540
Db	481	ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCCGATGATAGAGATAGAAA	540
Qy	541	TCTACTTTTGGGAAAGAGTGGTGGCTTTTGTCTGTGAAGAGGATTTCTTCTCAGGA	600
Db	541	TCTACTTTTGGGAAAGAGTGGTGGCTTTTGTCTGTGAAGAGGATTTCTTCTCAGGA	600
Qy	601	TCCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGTCTTATGCCAGACTCAGCTTTTTC	660
Db	601	TCCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGTCTTATGCCAGACTCAGCTTTTTC	660
Qy	661	AATGAACCTCATCAGCAGAGATGAAGACTTCTACTGTGACTTTTCTTGCCTGATGTTCCAA	720
Db	661	AATGAACCTCATCAGCAGAGATGAAGACTTCTACTGTGACTTTTCTTGCCTGATGTTCCAA	720
Qy	721	TACTTAGTAAATAGCCTTCAGAAAGAGGTTCAGGAGATCATTTGTGATGCTGTCAAA	780
Db	721	TACTTAGTAAATAGCCTTCAGAAAGAGGTTCAGGAGATCATTTGTGATGCTGTCAAA	780
Qy	781	ATTGACGAGAGTTTTAAACAGAGCTTCCAGTTGGCTCATTTGGATGAATTCGATT	840
Db	781	ATTGACGAGAGTTTTAAACAGAGCTTCCAGTTGGCTCATTTGGATGAATTCGATT	840
Qy	841	TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTTGTGAACTTGGATTTCTCA	900
Db	841	TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTTGTGAACTTGGATTTCTCA	900
Qy	901	AAGGTTTTTTCAGGAGAAATCTTTTGTATTTTATGAAACATTTCTTTTGAAGGAAA	960
Db	901	AAGGTTTTTTCAGGAGAAATCTTTTGTATTTTATGAAACATTTCTTTTGAAGGAAA	960
Qy	961	ACAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTATGGCAGAAACC	1020
Db	961	ACAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTATGGCAGAAACC	1020
Qy	1021	ACAGATAACGTCTTACCTTTGGATGAGATTTT 1053	

Db 1021 ACAGATAACGTCTTACCTTTGGATGAGATTTT 1053

RESULT 2

US-10-698-228-4
; Sequence 4, Application US/10698228
; Publication No. US2004007253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619W00P
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 4
; LENGTH: 1081
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-4

Query Match 100.0%; Score 1053; DB 19; Length 1081;
Best Local Similarity 100.0%; Pred. No. 9.4e-295;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATCAGATCATCTTCA	60
Db	20	ATGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATCAGATCATCTTCA	79
Qy	61	GACACCAACGAAAGTGAATAAAGTCAATGAAGACCACTCTTAAGAAAGAGTCTCGC	120
Db	80	GACACCAACGAAAGTGAATAAAGTCAATGAAGACCACTCTTAAGAAAGAGTCTCGC	139
Qy	121	CGGTTTGTCTATTTTCCAATCCAGTACCTCATATTTGGAAATGTATAAACAGGCACAG	180
Db	140	CGGTTTGTCTATTTTCCAATCCAGTACCTCATATTTGGAAATGTATAAACAGGCACAG	199
Qy	181	GCTTCTCTTCTGGACAGCAGAGAGTTCATTTATCAAAAGGATCTCCCTCACTGGAAACAAG	240
Db	200	GCTTCTCTTCTGGACAGCAGAGAGTTCATTTATCAAAAGGATCTCCCTCACTGGAAACAAG	259
Qy	241	CTTAAAGCAGATGAGAAGTACTTCTCTCATCATCTTTAGCCTTTTTCAGCCAGTGAT	300
Db	260	CTTAAAGCAGATGAGAAGTACTTCTCTCATCATCTTTAGCCTTTTTCAGCCAGTGAT	319
Qy	301	GGAAATGTAATGAAATTTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT	360
Db	320	GGAAATGTAATGAAATTTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT	379
Qy	361	CGCTGTTTCTATGGCTTTCAAATTTCTCATCAGAAATGTTTCACTCAGAGATGTACAGTTTG	420
Db	380	CGCTGTTTCTATGGCTTTCAAATTTCTCATCAGAAATGTTTCACTCAGAGATGTACAGTTTG	439
Qy	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATGAA	480
Db	440	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATGAA	499
Qy	481	ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCCGATGATAGCAGATAGAAA	540
Db	500	ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCCGATGATAGCAGATAGAAA	559
Qy	541	TCTACTTTTGGGAAAGAGTGGTGGCTTTTGTCTGTAGAGAGGATTTTCTTCTCAGGA	600
Db	560	TCTACTTTTGGGAAAGAGTGGTGGCTTTTGTCTGTAGAGAGGATTTTCTTCTCAGGA	619

```
QY 601 TCTTTGCTGCTATATTTCTGGCTAAAGAGAGAGGTTCTTATGCCAGGACTCACTTTTTC 660
Db 620 TCTTTGCTGCTATATTTCTGGCTAAAGAGAGAGGTTCTTATGCCAGGACTCACTTTTTC 679
QY 661 AATGAACCTCATCAGCAGAGATGAGAGCTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 680 AATGAACCTCATCAGCAGAGATGAGAGCTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 739
QY 721 TACTTAGTAAATAAGCTTTTAAAGAGAGAGGTTCAAGGAGATCACTTGTGATGCTGCMAA 780
Db 740 TACTTAGTAAATAAGCTTTTAAAGAGAGAGGTTCAAGGAGATCACTTGTGATGCTGCMAA 799
QY 781 ATTGAGCAGGAGTTTAAAGAGAGCTTTGCCAGTTGGCTCATTTGGAATGAATTCGATT 840
Db 800 ATTGAGCAGGAGTTTAAAGAGAGCTTTGCCAGTTGGCTCATTTGGAATGAATTCGATT 859
QY 841 TTGATGAACAGTACATTCAGTTGTTAGCTGACAGATTAATTGAGAACTTTGGAATTCCTCA 900
Db 860 TTGATGAACAGTACATTCAGTTGTTAGCTGACAGATTAATTGAGAACTTTGGAATTCCTCA 919
QY 901 AAGGTTTTTCAGGCAGAGAAATCCTTTTGATTTTATGGAACAATTTCTTTAGAGGAAA 960
Db 920 AAGGTTTTTCAGGCAGAGAAATCCTTTTGATTTTATGGAACAATTTCTTTAGAGGAAA 979
QY 961 ACAAAATTTCTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTATGCGCAGAAACC 1020
Db 980 ACAAAATTTCTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTATGCGCAGAAACC 1039
QY 1021 ACAGATAACGCTTTCACCTTGGATGCGAGATTTT 1053
Db 1040 ACAGATAACGCTTTCACCTTGGATGCGAGATTTT 1072
```

RESULT 3

```
US-10-698-228-2
; Sequence 2, Application US/10698228
; Publication No. US20040072253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-2
```

```
Query Match 99.8%; Score 1051.4; DB 19; Length 1053;
Best Local Similarity 99.9%; Pred. No. 2.7e-294;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 ATGGCGGACCCGGAAGCGCGGAGCTGGATCAGATGAGAGATCATCTTCA 60
Db 1 ATGGCGGACCCGGAAGCGCGGAGCTGGATCAGATGAGAGATCATCTTCA 60
QY 61 GACACCAACGAAAGTGAATAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAACGAAAGTGAATAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
```

```
QY 121 CGGTTTGTCTATCTTTTCAATCCAGTACCTTGATATTTGGAAAAATGTATAAACAGGACAG 180
Db 121 CGGTTTGTCTATCTTTTCAATCCAGTACCTTGATATTTGGAAAAATGTATAAACAGGACAG 180
QY 181 GCTTCTTCTTGGCAGCAGAGAGGTTGACTTATCAAAGGATCTCCCTCACTGGAAACAAG 240
Db 181 GCTTCTTCTTGGCAGCAGAGAGGTTGACTTATCAAAGGATCTCCCTCACTGGAAACAAG 240
QY 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCACATCTTTAGCCTTTTGTGACCCAGTGTAT 300
Db 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCACATCTTTAGCCTTTTGTGACCCAGTGTAT 300
QY 301 GGAATTTGTAATGAAAAATTTGGTGAGCGCTTTTAGTCAGAGAGTGCGAGTTCCAGAGGCT 360
Db 301 GGAATTTGTAATGAAAAATTTGGTGAGCGCTTTTAGTCAGAGAGTGCGAGTTCCAGAGGCT 360
QY 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCTGAGAAATGTTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCTGAGAAATGTTTCACTCAGAGATGTACAGTTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAATGCAATTTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAATGCAATTTGAA 480
QY 481 ACCATGCCCTATGTTAAAGAAAAAGCAGATTTGGGCTTTGCCGATGATAGCAGATAGAAAA 540
Db 481 ACCATGCCCTATGTTAAAGAAAAAGCAGATTTGGGCTTTGCCGATGATAGCAGATAGAAAA 540
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAAAGGAGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAAAGGAGTTTCTTCTCAGGA 600
QY 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
Db 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
QY 661 AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 661 AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
QY 721 TACTTAGTAAATAAGCTTTCAGAAAGAGGCTCAGGAGATCATTTGTTGATGCTGTCAAA 780
Db 721 TACTTAGTAAATAAGCTTTCAGAAAGAGGCTCAGGAGATCATTTGTTGATGCTGTCAAA 780
QY 781 ATTGAGCAGGAGTTTAAAGAGAGCTTGCAGGCTTGCCTCATTTGGAATGAATTCGATT 840
Db 781 ATTGAGCAGGAGTTTAAAGAGAGCTTGCAGGCTTGCCTCATTTGGAATGAATTCGATT 840
QY 841 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTAATTGTTGGAACCTTGGATTTCTCA 900
Db 841 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTAATTGTTGGAACCTTGGATTTCTCA 900
QY 901 AAGGTTTTTTCAGGCAGAGAAATCCTTTTGATTTTATGGAACAATTTCTTTAGAAAGGAAA 960
Db 901 AAGGTTTTTTCAGGCAGAGAAATCCTTTTGATTTTATGGAACAATTTCTTTAGAAAGGAAA 960
QY 961 ACAAAATTTCTTTCAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTATGCGCAGAAACC 1020
Db 961 ACAAAATTTCTTTCAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTATGCGCAGAAACC 1020
QY 1021 ACAGATAACGCTTTCACCTTGGATGCGAGATTTT 1053
Db 1021 ACAGATAACGCTTTCACCTTGGATGCGAGATTTT 1053
```

RESULT 4

```
US-10-172-118-71
; Sequence 71, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
```

APPLICANT: Mao, Mao
APPLICANT: Roberts, Chris
APPLICANT: Van 't Veer, Laura
APPLICANT: Van de Vijver, Marc
APPLICANT: Bernards, Rene
TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
FILE REFERENCE: 9301-175-999
CURRENT APPLICATION NUMBER: US/10/172,118
CURRENT FILING DATE: 2002-06-14
PRIOR APPLICATION NUMBER: 60/380,770
PRIOR FILING DATE: 2002-05-14
NUMBER OF SEQ ID NOS: 2699
SEQ ID NO 71
LENGTH: 4955
TYPE: DNA
ORGANISM: Homo sapiens
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: AB036063
DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-71

Query Match 99.8%; Score 1051.4; DB 18; Length 4955;
Best Local Similarity 99.9%; Pred. No. 6.8e-294;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAGCGCGGAGCGCGGGCTGGATCAGGATGAGAGATCATCTTCA 60
DB 245 ATGGGCGACCCGGAAGCGCGGAGCGCGGGCTGGATCAGGATGAGAGATCATCTTCA 304
QY 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
DB 305 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 364
QY 121 CGGTTTGTGATCTTTCGAATCCAGTACCTGATATTTGGAAATGATATAACAGGCACAG 180
DB 365 CGGTTTGTGATCTTTCGAATCCAGTACCTGATATTTGGAAATGATATAACAGGCACAG 424
QY 181 GCTTCTCTCTGGACACGAGAGGTGACTTATCAAGAGTCTCCTCACTCGAACAAG 240
DB 425 GCTTCTCTCTGGACACGAGAGGTGACTTATCAAGAGTCTCCTCACTCGAACAAG 484
QY 241 CTTAAAGCAGATGAGAAGTACTTCTATCTCTCACATCTTAGCCCTTTTTCAGCCAGTGAT 300
DB 485 CTTAAAGCAGATGAGAAGTACTTCTATCTCTCACATCTTAGCCCTTTTTCAGCCAGTGAT 544
QY 301 GGAATTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
DB 545 GGAATTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 604
QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCTATCGAGAATGTTTCACTCAGAGATGTACAGTTG 420
DB 605 CGCTGTTTCTATGGCTTTCAAAATCTCTATCGAGAATGTTTCACTCAGAGATGTACAGTTG 664
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTGAA 480
DB 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTGAA 724
QY 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGCCCTTTCGATGGATAGCAGATAGAAA 540
DB 725 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGCCCTTTCGATGGATAGCAGATAGAAA 784
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTTCTGCTGTAGAGAGGTTTCTTCTCAGGA 600
DB 785 TCTACTTTTGGGAAAGAGTGGTGGCTTTTCTGCTGTAGAGAGGTTTCTTCTCAGGA 844
QY 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGTCTTATGCGAGGACTCACTTTTTC 660
DB 845 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGTTTATGCGAGGACTCACTTTTTC 904
QY 661 AATGAACATCAGCAGAGATGAAGGACTTCACTGTGACTTTTCTGCTGCTGATGTTCCAA 720
DB 905 AATGAACATCAGCAGAGATGAAGGACTTCACTGTGACTTTTCTGCTGCTGATGTTCCAA 964

QY 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGCTCAGGAGAGATCATTTGTGATGCTGTCAA 780
DB 965 TACTTAGTAAATAAGCCTTCAGAAAGAGGCTCAGGAGAGATCATTTGTGATGCTGTCAA 1024
QY 781 ATTGAGCAGAGTATTTTAAACAGAGCCTTCCAGTTCGCTCATTTGGAATGAATTGCATT 840
DB 1025 ATTGAGCAGAGTATTTTAAACAGAGCCTTCCAGTTCGCTCATTTGGAATGAATTGCATT 1084
QY 841 TTGATGAACACAGTACATTTGAGTTTGTAGCTGACAGATTACTTTGGAACTTTGGATTCTCA 900
DB 1085 TTGATGAACACAGTACATTTGAGTTTGTAGCTGACAGATTACTTTGGAACTTTGGATTCTCA 1144
QY 901 AAGGTTTTTTCAGGCAGAAAAATCCTTTTGAATTTTATGAAAAACATTTCTTTAGAGAAAA 960
DB 1145 AAGGTTTTTTCAGGCAGAAAAATCCTTTTGAATTTTATGAAAAACATTTCTTTAGAGAAAA 1204
QY 961 ACAAATTTCTTTCAGAAACAGATTTTCAGATATCAGCGTTTTTGAGTTATGGCAGAAACC 1020
DB 1205 ACAAATTTCTTTCAGAAACAGATTTTCAGATATCAGCGTTTTTGAGTTATGGCAGAAACC 1264
QY 1021 ACAGATAACGCTTTCACCTTTGGATGCGAGATTTT 1053
DB 1265 ACAGATAACGCTTTCACCTTTGGATGCGAGATTTT 1297

RESULT 5
US-10-342-887-71
; Sequence 71, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 71
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-71

Query Match 99.8%; Score 1051.4; DB 19; Length 4955;
Best Local Similarity 99.9%; Pred. No. 6.8e-294;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAGCGCGGAGCGCGGGCTGGATCAGGATGAGAGATCATCTTCA 60
DB 245 ATGGGCGACCCGGAAGCGCGGAGCGCGGGCTGGATCAGGATGAGAGATCATCTTCA 304
QY 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
DB 305 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 364
QY 121 CGGTTTGTGATCTTTCGAATCCAGTACCTGATATTTGGAAATGATATAACAGGCACAG 180
DB 365 CGGTTTGTGATCTTTCGAATCCAGTACCTGATATTTGGAAATGATATAACAGGCACAG 424
QY 181 GCTTCTCTCTGGACACGAGAGGTGACTTATCAAGAGTCTCCTCACTCGAACAAG 240

QY 841 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTACTTGTGAACTTGGATTCTCA 900
Db 1085 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTACTTGTGAACTTGGATTCTCA 1144
QY 901 AGGTTTTCAGGCGAGAAATCCTTTTGTATTTATGGAACAACTTTTGAAGGAAAA 960
Db 1145 AAGGTTTTCAGGCGAGAAATCCTTTTGTATTTATGGAACAACTTTTGAAGGAAAA 1204
QY 961 ACAAATTTCTTTGAGAAACAGAGTTTCAGAGTATCAGGGTTTTCAGGTTATGCGAGAAACC 1020
Db 1205 ACAAATTTCTTTGAGAAACAGAGTTTCAGAGTATCAGGGTTTTCAGGTTATGCGAGAAACC 1264
QY 1021 ACAGATAACGTTCTTACCTTTGGATGAGAGTTT 1053
Db 1265 ACAGATAACGTTCTTACCTTTGGATGAGAGTTT 1297

RESULT 7

US-10-220-335-514/c
; Sequence 514, Application US/10220335
; Publication No. US20050175607A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-041
; CURRENT APPLICATION NUMBER: US/10/220,335
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 09/664,641
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/616,807
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 09/597,707
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/577,409
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: 09/515,126
; PRIOR FILING DATE: 2000-02-28
; NUMBER OF SEQ ID NOS: 688
; SOFTWARE: Custom
; SEQ ID NO 514
; LENGTH: 2482
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-220-335-514

Query Match 55.1%; Score 580; DB 24; Length 2482;
Best Local Similarity 75.1%; Pred. No. 4.5e-157;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;
QY 89 ATGAAGAGCCACTCCTAAGAAAGAGTTCTCGCGGTTTGTCTATCTTCCAAATCCAGTACC 148
Db 2086 AGATGAGCGCGTGCCTGAGAGAAACCCCGCGGCTTGTCTATCTTCCCATCGAGTACC 2027
QY 149 CTGATATTTGAAATGTATAAAGCAGGACAGGCTTCTCTGAGCAGAGAGAGGTTG 208
Db 2026 ATGATATCTGGCAGATGTATAAGAGCGAGAGGCTTCTTTGGACCGCGAGAGGTTG 1967
QY 209 ACTTATCAAGAGATCTCCCTCACTGGAAACAGCTTAAAGCAGATGAGAAGTACTTCACT 268
Db 1966 ACCTCTCAAGAGCAATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTATAT 1907
QY 269 CTCATCTTAGCCTTTTTCAGCCAGTATGGAATGTAATGAAATTTGGTGAGC 328
Db 1906 CCCATGTTCTGGCTTTCTTTGCGAGCAAGCGATGTAATGAAATTTGGTGAGC 1847
QY 329 GCTTTAGTCAGGAGGTCAGGTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAATCTCA 388
Db 1846 GATTTAGCCAAAGTTTCAGATTACAGAACCCGCTGTTTCTATGGCTTCAAATTCGCA 1787
QY 389 TCAGAGATGTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAGATCCCA 448
Db 1786 TGGAAACATACATCTTGAATGTATAGTCTTCTTATGACACTTACATAAAGATCCCA 1727

QY 449 AGAAAGGGGAATTTTATTAATGCAATGAAACCAATGCCCTATGTTAAAGAAAAAGCAG 508
Db 1726 AAGAAAGGGGAATTTCTCTCAATGCCATTTGAAACGATGCTTGTCAAGAAAGAGCAG 1667
QY 509 ATTGGGCTTTCGATGATGATGAGATGAGAAATCTACTTTTGGGGAAGAGTGTGGCT 568
Db 1666 ACTGGGCTTTCGCTGATTTGGGGAAGAGGCTACTATGATGATGATGATGATGATGAT 1607
QY 569 TTGCTGCTGTAGAGGAGTTTCTTCTCAGATCTTTTGTGCTGCTATATTTCTGGCTAAAGA 628
Db 1606 TTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1547
QY 629 AGAGAGGCTTATGCCAGGACTCACTTTTCCAAATGAACTCATCAGCAGAGATGAAGAC 688
Db 1546 AACGAGGACTGATGCTGGCTCACAATTTCTAATGAACCTTATAGCAGAGATGAGGTT 1487
QY 689 TTCACTGTGATTTGCTTGGCTGATGTTTCAATCTTCAATCTTCAATCTTCAATCTTCA 748
Db 1486 TACACTGTGATTTTGTCTTGGCTGATGTTTCAACACCTGGTACACAAACCATCGGAGGAG 1427
QY 749 GGGTCAGGAGATCATTTGATGCTGCTCAAAATTTGAGCAGGAGTTTAAACAGAGCCT 808
Db 1426 GAGTAAGAGAAATTAATTAATCAATGCTGTTGATGAGAGAGGAGTTCTCTAGGCT 1367
QY 809 TGCAGTTGGCTCATTTGGAATGAAATTCATTTTGTATGAAACAGTACATTTGATG 868
Db 1366 TGCTGTGAAGCTCATTTGGGATGAAATTTGCACTCTAATGAAGCAATACATTTGATG 1307
QY 869 CTGACAGATTACTTGTGGAATCTTGGATCTCAAGGTTTTCAGGCGAGAAATCCTTTTG 928
Db 1306 CAGACAGACTTATGCTGGAACCTGGTTTATAGCAAGGTTTTCAGAGTAGAGAACCCATTTG 1247
QY 929 ATTTTATGGAACATTTCTTTAGAAGGAAACAAATTTCTTTCAGAGAAACGAGTTTCAG 988
Db 1246 ACTTTATGAGAAATATTTTCACTGGAAGAAAGACTAACTTTTCAGAGAGAGATGAGCG 1187
QY 989 AGTATCAGCGTTTTCAGATTATGCGAGAAACCAACAGATAACGCTTTCACCTTGAATG 1048
Db 1186 AGTATCAGAGATGGGAGTGTGTCAAGTCCAACAGAGAAATCTTTTACCTTGGATGCTG 1127
QY 1049 ATTT 1052
Db 1126 ACTT 1123

RESULT 8

US-09-954-456-724
; Sequence 724, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840

;
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 724
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-724

Query Match 55.1%; Score 580; DB 9; Length 2500;
Best Local Similarity 75.1%; Pred. No. 4.6e-157;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

QY 89 ATGAAGAGCCACTCTAAGAAAGAGTTCTCGCGGTTGTCTATCTTCCATCCAGTACC 148
DB 397 AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTGTCTATCTTCCCATCCAGTACC 456
QY 149 CTGATATTTGGAAATGTATAACAGCGCACAGGCTTCTCTCGACAGCAGAGAGGTTG 208
DB 457 ATGATATCTGGCAGATGTATAGAACGACAGAGGCTTCTTTTGGACCGCGAGAGGTTG 516
QY 209 ACTTATCAAGAGTCTCCCTCACTGGAAACAAGCTTAAAGCAGATGAGAAGTATCTCATCT 268
DB 517 ACCTCTCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTATAT 576
QY 269 CTCACATCTTAGCCCTTTTTCGACGATGGAATGTAAATGAAATTTGGTGAGC 328
DB 577 CCCATGTTCTGGCTTTCTTTGACGACGAGCGATGCGATAGTAATGAAACTTGGTGAGC 636
QY 329 GCTTTAGTCAGAGGTCAGGTTCCAGAGGCTCGCTGTTTCTATGCTTCCAAATCTCA 388
DB 637 GATTTAGCCAAAGAGTTCAGATTACAGAACCCGCTGTTTCTATGCTTCCAAATGCCA 696
QY 389 TCAGAGATGTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAGATCCCA 448
DB 697 TGGAAACATACATCTGAAAGTATAGTCTTCTTATGACATTTACATAAAGATCCCA 756
QY 449 AGAAAGGGAATTTTATTAATGCAATGAAACCACTGCTATGTTAAAGAAAGAGCAG 508
DB 757 AAGAAAGGGAATTTCTCTTCAATGCCATTGAAACGATGCTTGTGTCAAGAGAGGAG 816
QY 509 ATTGGGCTTGGATGGATAGCAGATAGAAATCTTACTTTTGGGAAAGAGTGCTGCT 568
DB 817 ACTGGGCTTGGCTGGATTTGGGACAAAGAGGCTACCTATGTTGAACTGTTGTAGCT 876
QY 569 TTGCTGCTGTAGAGGAGTTTCTCTCAGGATCTTTTGTGCTATATTTCTGCTAAAGA 628
DB 877 TTGCTGCTGTAGAGGCAATTTCTTTCCGGTTCTTTTGGCTGATATTTCTGCTCAAGA 936
QY 629 AGAGAGGCTTATGCCAGGACTCATTCTTCCAAATGAACCTCATCAGCAGAGATGAAGGAC 688
DB 937 AACGAGGACTGATGCTGCTGCTCACATTTCTTAATGAACCTTATAGCAGATGAGGTT 996
QY 689 TTCACGTGACATTTGCTGCTGATGTTCCAAATCTTAGTAAATAGCCTTCAGAGAAA 748
DB 997 TACACTGTGATTTGCTGCTGATGTTCAAAACACCTGCTACAAACCACTCGAGGAGA 1056
QY 749 GGTCAAGGAGATCATTTGATGCTGTCAAAATTCAGCAGGAGTTTAAACAGAGCCT 808
DB 1057 GAGTAGAGAAATAATATCAATGCTGTTGGATAGAACAGGAGTTCTCCTCAGTGGCCT 1116
QY 809 TGCCAGTTGGCCTCATTTGGAATGAATGCAATTTTGTATGAAACAGTACATTTGATG 868
DB 1117 TGCTGTGAGCTCATTTGGATGAATTTGACCTTATAGAGCAATACATTTGATTTGTG 1176
QY 869 CTGACAGATTAATTGTGGAACTTGGATTTCTCAAAGGTTTTCAGGAGAGAAATCTTTTG 928
DB 1177 CAGACAGACTTATGCTGGAACCTGGGTTTATAGCAAGGTTTTCAGAGTAGAGAACCCATTG 1236
QY 929 ATTTATGAAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAACGAGTTTCAG 988

DB 1237 ACTTTATGAGAAATATTTTCACTGGAAGAAAGACTAATCTTTTGAAGAGAGTAGGCG 1296
QY 989 AGTATCAGCGTTTTCAGTTATGCGAGAAACCCAGAGATAACGTCTTCCACCTTGGATGAC 1048
DB 1297 AGTATCAGAGATGGGAGTGTATGTCAAAGTCCACAGAGAAATCTTTTACCTTTGGATGCTG 1356
QY 1049 ATTT 1052
DB 1357 ACTT 1360

RESULT 9

US-09-954-456-1169
; Sequence 1169, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cancer
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1169
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-1169

Query Match 55.1%; Score 580; DB 9; Length 2500;
Best Local Similarity 75.1%; Pred. No. 4.6e-157;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

QY 89 ATGAAGAGCCACTCTAAGAAAGAGTTCTCGCGGTTGTCTATCTTCCATCCAGTACC 148
DB 397 AGGATGAGCGCTGCTGAGAGAAACCCCGCGCTTGTCTATCTTCCCATCCAGTACC 456
QY 149 CTGATATTTGGAAATGTATAACAGCGCACAGGCTTCTCTCGACAGCAGAGAGGTTG 208
DB 457 ATGATATCTGGCAGATGTATAGAACGACAGAGGCTTCTTTTGGACCGCGAGAGGTTG 516
QY 209 ACTTATCAAGAGTCTCCCTCACTGGAAACAAGCTTAAAGCAGATGAGAAGTATCTCATCT 268
DB 517 ACCTCTCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTATAT 576
QY 269 CTCACATCTTAGCCCTTTTTCGACGATGGAATGTAAATGAAATTTGGTGAGC 328
DB 577 CCCATGTTCTGGCTTTCTTTGACGACGAGCGATGCGATAGTAATGAAACTTGGTGAGC 636
QY 329 GCTTTAGTCAGAGGTCAGGTTCCAGAGGCTCGCTGTTTCTATGCTTCCAAATCTCA 388
DB 637 GATTTAGCCAAAGAGTTCAGATTACAGAACCCGCTGTTTCTATGCTTCCAAATGCCA 696

389 TCAGAAATGTTCACTCAGAGATGTACAGTTTGTGTATAGACACTTATACATCAGAGATCCCA 448
Db TCGAATAACATATCTGAAATGTAGTCTTATTGACACTTATACATAAAGATCCCA 756
QY AGAAAGGAAATTTTATTATGCAATGAAACCAATGCCCTATGTTAAGAAAAAGCAG 508
Db AAGAAAGGAAATTTCTCTCAATGCAATGAAACCAATGCCCTATGTTAAGAAAGCAG 816
QY ATTGGGCTTGGATGATGAGATAGAAATCTACTTTTGGGAAAGAGTGGTGGCT 568
Db ACTGGGCTTGGATGATGAGATAGAAATCTACTTTTGGGAAAGAGTGGTGGCT 876
QY TTGCTGCTGTAAGAGAGTTTCTTCTCAGGATCTTTTCTGCTATATTTCTGCTAAAGA 628
Db TTGCTGCTGTAAGAGAGTTTCTTCTTCCGTTCTTTTGGGTCGATTTCTGCTCAAGA 936
QY AGAGAGTCTTATGCAAGACTCATTTTTCATGAACTCATCAGCAGATGAAGAC 688
Db AACGAGGACTGATGCTGCTGCTCACAATTTCTAATGAACCTTATAGCAGATGAGGTT 996
QY TTCACTGCTGACCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 748
Db TACACTGTGATTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1056
QY GGTTCAGGAGATCATTTGTTGATGCTGCTCAAAATTCAGCAGAGTTTAAACAGAGCCT 808
Db GGTTCAGGAGATCATTTGTTGATGCTGCTCAAAATTCAGCAGAGTTTAAACAGAGCCT 1116
QY TGCCAGTTGGCTCATTTGGAATGAATGCAATTTTGAAGAAACAGTACATGAGTTGTAG 868
Db TGCCAGTTGGCTCATTTGGAATGAATGCAATTTTGAAGAAACAGTACATGAGTTGTAG 1176
QY CTGACAGATTAATTTGGAATGCAATTTTGAAGAAACAGTACATGAGTTGTAG 928
Db CTGACAGATTAATTTGGAATGCAATTTTGAAGAAACAGTACATGAGTTGTAG 1236
QY ATTTTATGAAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACAGTTTTCAG 988
Db ACTTTATGGAATATTTCTGGAAGGAAAGACTAACTTTCTTTGAGAGAGTAGGCG 1296
QY AGTATCAGGTTTGCAGTTATGGCAGAAACCAAGATACGCTTTTCACTTGGATGCGAG 1048
Db AGTATCAGGTTTGCAGTTATGGCAGAAACCAAGATACGCTTTTCACTTGGATGCGAG 1356
QY ATTT 1052
Db ACTT 1360

RESULT 10
US-09-954-456-1827
; Sequence 1827, Application US/09954456
; Patent No. US2002115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26

; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1827
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-1827

Query Match 55.1%; Score 580; DB 9; Length 2500;
Best Local Similarity 75.1%; Pred. No. 4.6e-157;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

QY 89 ATCAAGAGCCACTCTAAGAAAGAGTTCTCGCCGGTTTGTCTATCTTTTCCAATCCAGTACC 148
Db 397 AGATGAGCCGCTGCTGAGAGAAACCCCGCCCTTTGTCTATCTTTCCCATCGATACC 456
QY 149 CTGATATTTGAAAAATGTATAAAGCAGGCACAGGCTTCTCTCGACAGCAGAAAGAGTTG 208
Db 457 ATGATATCGCAGATGTATAAGAGGCAGAGGCTTCTTTTGGACCCGCGAGAGTTG 516
QY 209 ACTTATCAAGAGATCTCCCTCAGTGGAAACAGCTTAAAGCAGATGAGAGTACTTCATCT 268
Db 517 ACCTCTCCAGGACATTCAGCACTCTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT 576
QY 269 CTCACATCTTAGCCTTTTTCAGCCAGTGAATTTGTAATGAAATTTGGTGGAGC 328
Db 577 CCATGTTCTGGCTTCTTTTCAGCAAGCAGATGATGAATGAAATCTTTGGTGGAGC 636
QY 329 GCTTTAGTCAGAGAGTGCAGGTTCCAGAGGCTCGCTGTTCTATGGCTTTCAAAATTTCTCA 388
Db 637 GATTTAGCAAGAAGTTCAGATTACAGAGCCGCTGTTCTATGGCTTTCCAAATTTGCCA 696
QY 389 TCAGAAATGTTCTCAGAGATGTACAGTTTGTGTAGACACTTACATCAGAGATGCCA 448
Db 697 TGAAGAAACATACATCTCGAAATGTATAGTCTTCTTATTGACACTTACATAAAGATGCCA 756
QY 449 AGAAAGGAAATTTTATTAAATGCAATTTGAAACCAATGCCCTATGTTAAGAAAAAGCAG 508
Db 757 AAGAAAGGAAATTTCTCTCAATGCCAATGAAACGATGCTTGTGTCAGAAAGAGGAG 816
QY 509 ATTGGGCTTGGATGATAGAGATAGAAATCTACTTTTGGGAAAGAGTGGTGGCT 568
Db 817 ACTGGGCTTGGCTGATTTGGGACAAAGAGGCTACCTATGGTGAACGTTGTAGCT 876
QY 569 TTGCTGCTGTAAGAGAGTTTCTTCTCAGAGATCTTTTGTGCTATATTTCTGGCTAAAGA 628
Db 877 TTGCTGCTGTAAGAGAGTTTCTTCTCCGTTCTTTTTCGCTCGATATTTCTGGCTCAAGA 936
QY 629 AGAGAGTCTTATGCCAGGACTCATTTTTCATGAACTCATCAGCAGATGAAGAC 688
Db 937 AACGAGGACTGATGCTGGCTCACAATTTTCTAATGAACCTTATTAGCAGATGAGGTT 996
QY 689 TTCACTGTGACTTTTGGCTGCTGATGTTTCAATCTTCTAGTAAATAGGCTTTTCAAGAGAA 748
Db 997 TACACTGTGATTTTGGCTGCTGATGTTTCAACCTGTTACACAAACCATCGGAGGAGA 1056
QY 749 GGCTCAGGAGATCATTTGTTGATGCTCTCAAAATTTGAGCAGGAGTTTAAACAGAGCCT 808
Db 1057 GAGTAAGAGAAATAATTTATCAATGCTGTTCGGATAGAAACAGAGTTCTCTCACTGAGGCT 1116
QY 809 TGGCAGTTGGCCTCATTTGGAATGAAATGCAATTTTGAAGAAACAGTACATGAGTTGTAG 868
Db 1117 TGCCTGTGAAGCTCATTTGGGATGAAATTTGCACTCTTAATGAAGCAATACATTTGAGTTGTGG 1176
QY 869 CTGACAGATTAATTTGTTGGAATTTGATTTCTCAAGGTTTTCAGGAGAGAAATCTTTTG 928

Db 1177 CAGACAGACTTATGCTGGAACCTGGGTTTGTAGCAAGGTTTTCAGATAGAGAACCCATTG 1236
QY 929 ATTTTATGAAACATTTCTTTAGAGGAAACAAATTTCTTTTGAAACGAGTTTCAG 988
Db 1237 ACTTTATGAGAAATATTTCACTGGAAGGAAAGACTAACTTTCTTTTGAGAGAGTAGGCG 1296
QY 989 AGTATCAGCGCTTTTGCAGTTATGGCAGAAACACACAGATAACGCTTTCACCTTCGATGCGAG 1048
Db 1297 AGTATCAGAGGATGGGAGTGATGTCAGTCCACACAGAGAAATTTTACCTTGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 11

US-10-641-643-1370
; Sequence 1370, Application US/10641643
; Publication No. US20040077003A1
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; Susan G. Stuart
; Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
; GENE EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/641,643
FILING DATE: 14-Aug-2003
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.

REGISTRATION NUMBER: 37,071

REFERENCE/DOCKET NUMBER: PA-0001 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 855-0555

TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 1370:

SEQUENCE CHARACTERISTICS:

LENGTH: 2500 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GENBANK

CLONE: g36154

SEQUENCE DESCRIPTION: SEQ ID NO: 1370 :

US-10-641-643-1370

Query Match 55.1%; Score 580; DB 19; Length 2500;

Best Local Similarity 75.1%; Pred. No. 4.6e-157; Indels 0; Gaps 0;

Matches 724; Conservative 0; Mismatches 240;

QY 89 ATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGCGGGTTTGTCTATCTTTTCAATCCAGTACC 148

Db 397 AGGATGAGCGCGTCTGTAGAGAAACCCCGCGGCTTGTCTATCTTTCCCATCGAGTACC 456

RESULT 12

US-10-733-878-458

; Sequence 458, Application US/10733878

; Publication No. US20040224408A1

; GENERAL INFORMATION:

; APPLICANT: Jean-Philippe Girard

; APPLICANT: Francois Amalric

; APPLICANT: Myriam Roussigne

Query Match	55.1%	Score 580;	DB 21;	Length 2500;
Best Local Similarity	75.1%	Pred. No. 4.6e-157;		
Matches 724; Conservative	0;	Mismatches 240;	Indels 0;	Gaps 0;
QY	89	ATCAAGAGCCACCTCCTAAAGAAAGAGTTCTCGCGGTTTGTTCATCTTCCAAATCCAGTACC	148	
DB	397	AGATGAGCGCTGCTGAGAGAAACCCCGCGCTTTGTTCATCTTCCCATCGAGTACC	456	
QY	149	CTGATATTTTGGAAAAATGTTATPAAACAGGCACAGGCTTCCTTCTGACACAGCAGAAGAGTTG	208	
DB	457	ATGATATCTGGCAGATGTTATPAGNAGCAGAGGCTTCCTTTTGGACCGCGAGAGGTTG	516	
QY	209	ACTTATCAAAGGATTCCTCTCATCTGGAACAAAGCTTAAAGCAGATGAGAAGTACTTCATCT	268	
DB	517	ACCTCTCCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTTATAT	576	
QY	269	CTCACATCTTAGCCCTTTTTCGACCCAGTGTGNAATTTGAATGAAAAATTTGCTGGAGC	328	
DB	577	CCCATGTTCTGGGCTTCTTTTTCGACAAAGCATGGCATAGTAAATGAAAACTTGGTGGAGC	636	
QY	329	GCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCTCGCTGTTCTATGGCTTTTCAAATTTCTCA	388	
DB	637	GATTTAGCCNAGAGTTTCAGATTACAGAGCCGCTGTTCTATGGCTTCCAAATTTGCCA	696	
QY	389	TCGAGAATGTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAGATCCCA	448	
DB	697	TGAAAAACATACATTTCTGAAATGTATAGTCTTCTTATTGACACTTACATAAAAGATCCCA	756	
QY	449	AGAAAAGGNAATTTTATTTTAAATGCAATTTGAAACCATGCCCTATGTTAAGAAAAAGGAG	508	
DB	757	AAGAAAAGGGAATTTCTCTTCAAATGCCATTTGAAACGATGCCCTTGTGTCAAGAAAGGCGAG	816	
QY	509	ATTGGGCTTTGGCATGGATAGCAGATAGAAAACTCATTTTGGGGAAGAGTGTGTGGCCT	568	
DB	817	ACTGGGCTTTGCCCTGGATTTGGGGAANAAGAGGCTACCTATGTTGAACGTGTTAGCCT	876	
QY	569	TTGCTGTGTGAGAAGAGATTTTCTTCTTCAGGATCTTTTGTGCTATATTTCTGGCTAAAGA	628	
DB	877	TTGCTGCAGTGAAGGCAATTTTCTTTTCCGGTCTTTTTCGTGATATTTCTGGCTCAAGA	936	
QY	629	AGAGAGTCTTATGCGCAGGACTCACTTTTTCGAATGAACCTCATCAGCAGAGATGAAGAC	688	
DB	937	AACGAGGACTGATGTCCTGGGCTCACATTTTCTAATGAACCTTATPAGCAGAGATGAGGTT	996	
QY	689	TTCACTGTGACTTTGCTTGCCTGATGTTCCAAATCTAGTAAATAAGCCCTTCAGAAANA	748	
DB	997	TACACTGTGATTTTGTCTGCTGATGTTCAACACCTGGTACAAAAACCATCGAGGAGA	1056	
QY	749	GGGTGAGGAGATCATTTGTTGATGCTGTCAAAATTTGAGCAGGAGTTTTTAACAGAAGCT	808	
DB	1057	GAGTAGAGNAATTAATCAATGCTGTTTGGATAGAACAGGAGTTCTCTCTGAGGCT	1116	
QY	809	TGCCAGTTGGCCTCATTTTGGAAATGAATTCATTTTGTATGAAACAGTACATTGAGTTGTAG	868	

Db 517 ACCTCTCCAGAGCAATTCAGCATCTGGGAATCCCTGAAACCCGAGGAGAGATATTTATAT 576
Qy 269 CTCACATCTTAGCCTTTTTCAGCCAGCATGGAATGTAAATGAAATTTTGGTGAGC 328
Db 577 CCCATGTTCTGGCTTTCTTTGACGACGATGCGCATAGTAATGAAACTTTGGTGAGC 636
Qy 329 GCTTTAGTCAGAGGTCAGAGTTCCAGAGGCTCGCTGTTTCTATGCTTTTCAAATTTCA 388
Db 637 GATTTAGCCAAGAAAGTTTCAGATTACAGAGCCGCTGTTTCTATGCTTCCAAATTTCCA 696
Qy 389 TCGAGATCTTCACACAGAGATCTACAGTTGCTGTATAGACATTCATCAGATCCCA 448
Db 697 TGAAGAACATACATCTGAAATGTATAGTCTTCTTATGACACTTACATAAAGATCCCA 756
Qy 449 AGAAAGGGAATTTTATTTAATGCAATGAAACCATGCTATGTTTAAAGAAAAGCAG 508
Db 757 AGAAGGGAATTTCTTCAATGCAATGAAACCATGCTATGTTTAAAGAAAAGCAG 816
Qy 509 ATTGGCCCTTGGATGAGATGACAGATAGAAAATCTACTTTTGGGAAAGAGTGGTGGCCT 568
Db 817 ACTGGCCCTTGGCTGATTTGGGACAAAGAGGCTACCTATGCTGAACGTTGTAGCCT 876
Qy 569 TTGCTGCTAGAGAGGTTTCTTCTCAGGATCTTTTGGTGTCTATTTCTGGCTAAAGA 628
Db 877 TTGCTGCTAGAGAGGCTTTTCTTCTTCCGGTCTTTTGGCTGATATTTCTGGCTCAAGA 936
Qy 629 AGAGAGGCTTATGCCAGGACTCACCTTTTCCAAATGAGAGGAGTCTTAAAGAGGCT 688
Db 937 AACGAGGACTGATGCTGGCCTCACATTTTCTAATGAACCTTATTTAGCAGAGATGAGGCT 996
Qy 689 TTCACTGTGACTTTGCTTGGCTGATGTTTCCAAATGAGTCTTAAAGAGGCTTCAAGAGAA 748
Db 997 TACACTGTGATTTGCTTGGCTGATGTTTCCAAATGAGTCTTAAAGAGGCTTCAAGAGAA 1056
Qy 749 GGGTCAGGAGATCAATGTTGATGCTGCTCAAAATGAGAGGAGTCTTAAAGAGGCT 808
Db 1057 GAGTAAGAGAAATATTAATCAATGCTGTTCCGATAGAACAGGAGTCTTCACTGAGGCT 1116
Qy 809 TGCCAGTTGGCTCTCATGGAATGAAATGCAATTTGATGAAACAGTACATTTGATG 868
Db 1117 TGCCGTGAGCTCATTTGGATGAAATGCAATTTGATGAAACAGTACATTTGATG 1176
Qy 869 CTGACAGATTAATCTGGAACCTTGGATCTCAAAAGTCTTCAAGGAGGAGTCTTCAAG 928
Db 1177 CAGACAGATTAATCTGGAACCTTGGATCTCAAAAGTCTTCAAGGAGGAGTCTTCAAG 1236
Qy 929 ATTTATGAAACATTTCTTTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 988
Db 1237 ACTTTATGAGAAATATTTCACTGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1296
Qy 989 AGTATCAGGCTTTGAGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1048
Db 1297 AGTATCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1356
Qy 1049 ATTT 1052
Db 1357 ACTT 1360

RESULT 14

US-10-843-641A-4196
; Sequence 4196, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals, Inc.
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; FILE OF INVENTION: Signature Gene Sets
; FILE REFERENCE: 689290-189
; CURRENT APPLICATION NUMBER: US/10/843,641A
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/873,367
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US/09/954,531

; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,832
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/964,824
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/09/967,768
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/09/968,007
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,708
; PRIOR FILING DATE: 2001-10-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 8447
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4196
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-843-641A-4196

Query Match 55.1%; Score 580; DB 22; Length 2500;
Best Local Similarity 75.1%; Pred. No. 4.6e-157;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

Qy 89 ATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGCCGGTTGTCTCATCTTTTCCAATCCAGTACC 148
Db 397 AGATGAGCCGCTGCTGAGAGAAACCCCGCCCTTTGTCTCATCTTCCCATCGAGTACC 456
Qy 149 CTGATATTTGAAATATATAACAGGCACAGGCTTCCCTCTGGACAGCAGAGAGGTTG 208
Db 457 ATGATATCTGCAGATGATATAAGAGGAGAGGCTTCTTTTGGACCCCGAGAGGTTG 516
Qy 209 ACTTATCAAAGAGTCTCCCTCAGTGGAAACAGCTTAAAGCAGATGAGAGTACTTCACT 268
Db 517 ACCTCTCAAAGAGCAATTCAGCACTCTGGGAATCCCTGAAACCCGAGAGAGATATTTATAT 576
Qy 269 CTCACATCTTAGCCTTTTTCAGCCAGTGTGAATGTAATGTAATGTAATGTAATGTAATG 328
Db 577 CCATGTTCTGGCTTTCTTTGACAGCAAGCGATGCGATGTAATGTAATGTAATGTAATG 636
Qy 329 GCTTTAGTCAGGAGGTCAGGTTCCAGAGGCTCGCTGTTCTATGGCTTTTCAAAATTTCTCA 388
Db 637 GATTTAGCCAAGAGTTTCAGATTACAGAGCCGCTGTTCTATGGCTTCCAAATTTGCCA 696
Qy 389 TCAGAGATGTTCTCAGAGATGTACAGTTTGTGTGATGACACTTACATCAGAGATCCCA 448
Db 697 TGAAGAACATACATCTCGAAATGTATAGTCTTCTTATTTGACACTTACATAAAGATCCCA 756
Qy 449 AGAAAGGGAATTTTATTTAATGCAATTCGAACCATGCGCTTGTGTCAAGAGAGAGGAG 508
Db 757 AGAAAGGGAATTTCTCTTCAATGCGCATGGAAGGCTTGTGTCAAGAGAGAGGAG 816
Qy 509 ATTGGCCCTTGGATGAGATGAGAAATCTACTTTTGGGAAAGAGTGGTGGCT 568
Db 817 ACTGGCCCTTGGCTGATTTGGGACAAAGAGGCTACTATGTTGAGAGGCTTTGATGCT 876
Qy 569 TTGCTGCTGAGAGGAGTCTTCTTCTCAGGATCTTTTGTGCTGATATTTCTGGCTAAAGA 628
Db 877 TTGCTGAGTGAAGGAGCAATTTCTTTTCCGGTCTTTTGGCTGATATTTCTGGCTCAAGA 936
Qy 629 AGAGAGGCTTATGCCAGGACTCACCTTTTCCAAATGAGAGGAGTCTTAAAGAGGCT 688
Db 937 AACGAGGACTGATGCTGGCCTCACATTTTCTAATGAACCTTATTTAGCAGAGATGAGGCT 996
Qy 689 TTCACTGTGACTTTGCTTGGCTGATGTTTCCAAATGAGTCTTAAAGAGGCTTCAAGAGAA 748
Db 997 TACACTGTGATTTGCTTGGCTGATGTTTCCAAATGAGTCTTAAAGAGGCTTCAAGAGAA 1056

QY 749 GGGTCAGGAGATCATTTGTTGATGCTGTCAAAATTTGAGCAGGAGTTTAAACAGAGCCT 808
Db 1057 GAGTAAGAGAAATTAATTAATCAATGCTGTTCCGATAGAAACAGGAGTTTCCCTACAGGCT 1116
QY 809 TGCCAGTTGGCCCTCAATTTGGAATGAAATTTGATGAAACAGTACATTTGAGTTTGTAG 868
Db 1117 TGCCGTGGAAGCTCATTTGGGATGAATTTGACCTCAATGAAGCAATACATTTGAGTTTGTG 1176
QY 869 CTGACAGATTACTTTGGAACTTTGGATTCTCAAGTTTTTTCAGGAGAAAATCCCTTTTG 928
Db 1177 CAGACAGACTTATTTGCTGGAATTTGGTTTATGCAAGTTTTTTCAGAGTAGAAGCAATTTG 1236
QY 929 ATTTTATGAAACAACTTTCTTTAGAGAGAAACAAATTTCTTTGAGAAACGAGTTTTCAG 988
Db 1237 ACTTTATGAGAAATATTTCTGAGAGAAAGCACTAACTTTCTTTGAGAGAGAGTAGGCG 1296
QY 989 AGTATCAGGCTTTTTCAGTTATTTGAGGAGAAACACAGATACGCTTTTCACTTTGAGTACAG 1048
Db 1297 AGTATCAGGAGTGGAGTGTCAAGTCCACAGAGAAATTTCTTTTACCTTTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360
RESULT 15
US-10-843-641A-4854
; Sequence 4854, Application US/10843641A
; Publication No. US2005006454A1
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals, Inc.
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; FILE OF INVENTION: Signature Gene Sets
; FILE REFERENCE: 689290-189
; CURRENT APPLICATION NUMBER: US/10/843,641A
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/873,367
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US/09/954,531
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,832
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/964,824
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/09/967,768
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/09/968,007
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,708
; PRIOR FILING DATE: 2001-10-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 8447
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4854
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-843-641A-4854

Query Match 55.1%; Score 580; DB 22; Length 2500;
Best Local Similarity 75.1%; Pred. No. 4.6e-157;
Matches 724; Conservative 0; Mismatches 240; Indels 0; Gaps 0;
QY 89 ATGAAGAGCACCTCCCTAAGAAAGAGTTCTCGCGGTTTGTCTCTTTCCAAATCCAGTACC 148
Db 397 AGATGAGCGCGTGTGAGAGAAACCCCGCGCTTTGTCACTCTCCCATCGATACC 456

QY 149 CTGATATTTGGAAAAATGTATAAAGCAGCAGGCTTCTCTCTTGACAGCAGAAAGGTTG 208
Db 457 ATGATATCTGGCAGATGTATAGAGGCGAGGCTTCTTTTGGACCGCCGAGGAGTTG 516
QY 209 ACTTATCAAGAGTCTCCCTCACTCGGAACAAAGCTTAAAGCAGATGAGAAGTACTTCACT 268
Db 517 ACCTCTCCAAGGACATTCAGCACTGGGAATCCCTGAAACCCGAGGAGAGATATTTATAT 576
QY 269 CTCACATCTTAGCCTTTTTCAGCCAGTGTGAATTTGTAATGAAATTTGGTGGAGC 328
Db 577 CCATGTTCTGGCTTTCTTTTCAGCAGCAAGCATGGCATAGTAATGAAATTTGGTGGAGC 636
QY 329 GCTTTAGTCAGGAGTGCAGGTTCCAGAGGCTGCTCTCTTTCTATGGCTTTTCAAAATTTCA 388
Db 637 GATTTAGCCAAAGTTTCAATTTACAGAGCCGCTGTTTCTATGGCTTTCCAAATTTGCA 696
QY 389 TCAGAAATGTTCACTCAGAGATGTACAGTTTGTGTAGACACTTACATCAGAGATCCCA 448
Db 697 TGAAGAACATACATTTCTGAAATGTATAGTCTTCTTATTGACACTTACATAAAGATCCCA 756
QY 449 AGAAAGGGAATTTTATTTAATGCAATTTGAAACCATGCCCTATGTTAAGAAAAGCAG 508
Db 757 AAGAAAGGGAATTTCTCTTCAATGCCATTTGAAACGATGCCCTTGTCTCAAGAAAGAGCAG 816
QY 509 ATTGGGCTTTCGATGGATAGCAGATAGAAAATCTACTTTTGGGGAAGAGTGGTGGCT 568
Db 817 ACTGGGCTTTCGCTGATTTGGGACAAAGAGGCTACTATGGTGAACGTTGTGTAGCT 876
QY 569 TTGCTGCTGTAGAGGAGTTTCTTCTCAGGATCTTTTGTGCTATATTTCTGGTAAAGA 628
Db 877 TTGCTGAGTGAAGGCAATTTCTTTTCCGCTTCTTTTGGCTCGATATTTCTGGCTCAAGA 936
QY 629 AGAGAGTCTTATGCGCAGGACTCACTTTTTCATGAACTCATCAGCAGAGATGAAGAC 688
Db 937 AACGAGGACTGATGCCCTGCGCTCACATTTTCTAATGAACCTTATTAGCAGAGATGAGGTT 996
QY 689 TTCACTGTGACTTTGCTTGCCTGATGTTCCAACTTAGTAAATGAAGCTTTCAGAAAGAA 748
Db 997 TACACTGTGATTTTGTCTGCTGATTTCAACACCTGGTACACAAACCATCGGAGGAGA 1056
QY 749 GGGTCAGGAGATCATTTGTTGATGCTGTCAAAATTTGAGCAGGAGTTTAAACAGAGCCT 808
Db 1057 GAGTAAGAGAAATTAATTAATCAATGCTGTTCCGATAGAAACAGGAGTTCTCTACTAGGCT 1116
QY 809 TGCCAGTTGGCCCTCAATTTGGAATGAAATTTGATGAAACAGTACATTTGAGTTTGTAG 868
Db 1117 TGCTGTGAAAGCTCATTTGGGATGAATTTGCACTCTTAATGAAGCAATACATTTGAGTTTGTG 1176
QY 869 CTGACAGATTACTTTGGAACTTTGGATTCTCAAGGTTTTTTCAGGAGAAAATCCCTTTG 928
Db 1177 CAGACAGACTTATGCTGGAATCTGGTTTTTTCAGAGTAGAAGCAATTTG 1236
QY 929 ATTTTATGAAACAACTTTCTTTTGAAGGAAAAAATAATTTCTTTGAGAAAACGAGTTTTCAG 988
Db 1237 ACTTTATGAGAAATATTTCTGAGAGAAAGCACTAACTTTCTTTGAGAGAGAGTAGGCG 1296
QY 989 AGTATCAGGCTTTTTCAGTTATTTGAGGAGAAACACAGATACGCTTTTCACTTTGAGTACAG 1048
Db 1297 AGTATCAGAGGATGGGAGTGTCAAGTCCACAGAGAAATTTCTTTTACCTTTGGATGCTG 1356
QY 1049 ATTT 1052
Db 1357 ACTT 1360

Search completed: October 30, 2005, 06:24:12
Job time : 679.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 30, 2005, 03:23:39 ; Search time 62.25 Seconds
(without alignments)
13084.616 Million cell updates/sec

Title: US-10-698-228-12

Perfect score: 1846

Sequence: 1 atggggaccggaaagcc.....tcaccttgatgcagatttt 1053

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 4211384

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+n2p.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US10698228/runat_26102005_100607_3971/app_query.fasta_1.2446
-DB=A_Geneseq_16Dec04 -QFMT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=ptc -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10698228 @CGN_1_140 @runat_26102005_100607_3971 -NCPU=6 -ICPU=3
-NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_16Dec04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1821	98.6	351	3 AAY84439	Aay84439 Amino aci
2	1821	98.6	351	4 AAB69050	Aab69050 Human rib
3	1817	98.4	351	4 AAB93497	Aab93497 Human pro
4	1454	78.8	389	4 AAB69051	Aab69051 Human R2
5	1454	78.8	389	6 ABU07433	Abu07433 Protein d
6	1454	78.8	389	7 ADJ68753	Adj68753 Human hea
7	1454	78.8	389	8 ADJ66564	Adj66564 Ribonucle
8	1454	78.8	389	8 ADK70435	Adk70435 Respirato
9	1454	78.8	389	8 ADN03789	Adn03789 Antipsori
10	1454	78.8	389	8 ADN04444	Adn04444 Antipsori

11	1454	78.8	389	8 ADO19226	Ado19226 Human PRO
12	1454	78.8	389	8 ADQ09272	Adq09272 Human REM
13	1454	78.8	389	8 ABM80258	Abm80258 Tumour-as
14	1454	78.8	413	3 AAB43902	Aab43902 Human can
15	1454	78.8	453	4 AAU28017	Aau28017 Human con
16	1454	78.8	453	4 ABG15287	Abg15287 Novel hum
17	1429	77.4	390	2 AAW41765	Aaw41765 Human rib
18	1425.5	77.2	386	8 ADO57309	Ado57309 Kidney de
19	1348	73.0	393	4 ABB64130	Abb64130 Drosophil
20	1242	67.3	430	8 ADN21469	Adn21469 Bacteri
21	1230	66.6	405	6 ABJ26187	Abj26187 Aspergill
22	1214	65.8	381	8 ADN23033	Adn23033 Bacteri
23	1192.5	64.6	381	6 ABJ25587	Abj25587 Aspergill
24	1192	64.6	413	5 ABP73700	Abp73700 Candida a
25	1176.5	63.7	321	8 ADS21332	Ads21332 Bacteri
26	1169.5	63.4	399	6 ABR53196	Abr53196 Protein s
27	1169.5	63.4	399	7 ADK63240	Adk63240 Disease t
28	1169.5	63.4	399	8 ADS43740	Ads43740 Bacteri
29	1161.5	62.9	391	8 ADS44135	Ads44135 Bacteri
30	1161.5	62.9	401	7 ADB70101	Adb70101 C. neofo
31	1132	61.3	352	3 AAG05697	Aag05697 Arabidops
32	1130	61.2	329	3 AAG51528	Aag51528 Arabidops
33	1130	61.2	329	3 AAG05699	Aag05699 Arabidops
34	1130	61.2	332	3 AAG51527	Aag51527 Arabidops
35	1130	61.2	332	3 AAG05698	Aag05698 Arabidops
36	1117	60.5	347	3 AAG14546	Aag14546 Arabidops
37	1112.5	60.3	330	3 AAG14548	Aag14548 Arabidops
38	1112.5	60.3	333	3 AAG14547	Aag14547 Arabidops
39	1099	59.5	314	3 AAG51529	Aag51529 Arabidops
40	1091	59.1	341	3 AAG41395	Aag41395 Arabidops
41	1091	59.1	363	3 AAG41394	Aag41394 Arabidops
42	1084.5	58.7	330	3 AAG41396	Aag41396 Arabidops
43	1072.5	58.1	274	2 AAW26418	Aaw26418 Swinepox
44	1072.5	58.1	274	4 AAB68239	Aab68239 Protein e
45	977.5	53.0	322	2 AAR72736	Aar72736 Plasmodiu

ALIGNMENTS

RESULT 1

AAY84439

ID AAY84439 standard; protein; 351 AA.

XX AAY84439;
AC AAY84439;

XX 25-JUL-2000 (first entry)

XX Amino acid sequence of a human RNA-associated protein.

XX Human; RNA-associated protein; cell proliferation; cancer; inflammation;
KW immune response; reproductive disorder; actinic keratosis;
KW atherosclerosis; arteriosclerosis; bursitis; cirrhosis; hepatitis;
KW mixed connective tissue disease; myelofibrosis; primary thrombocythemia;
KW paroxysmal nocturnal hemoglobinuria; polycythemia vera; psoriasis;
KW trauma.

XX Homo sapiens.

XX Key Location/Qualifiers
FH Key Modified-site 18
FT Modified-site 23

FT Modified-site 23 /note= "potential phosphorylation site"
FT Modified-site 29 /note= "potential glycosylation site"
FT Modified-site 38 /note= "potential phosphorylation site"
FT Modified-site 39 /note= "potential phosphorylation site"
FT Modified-site 65 /note= "potential phosphorylation site"
FT Modified-site 87 /note= "potential phosphorylation site"
FT Modified-site 87 /note= "potential phosphorylation site"

FT Modified-site 182 /note= "potential phosphorylation site"
 FT Modified-site 225 /note= "potential phosphorylation site"
 FT Modified-site 314 /note= "potential glycosylation site"
 FT Modified-site 329 /note= "potential phosphorylation site"
 XX
 PN WO200015799-A2.
 XX
 PD 23-MAR-2000.
 XX
 PF 17-SEP-1999; 99WO-US021688.
 XX
 PR 17-SEP-1998; 98US-00156039.
 PR 22-SEP-1998; 98US-00158720.
 PR 04-NOV-1998; 98US-00186815.
 PR 08-APR-1999; 99US-0128660P.
 XX
 PA (INCY-) INCYTE PHARM INC.
 XX
 PI Tang YT, Corley NC, Guegler KJ, Gorgone GA, Patterson C;
 PI Hillman JL, Baughn MR, Lal P, Azimzai Y, Yue H, Yang J;
 XX
 DR WPI: 2000-271437/23.
 DR N-PSDB; AAI12411.
 XX
 PR New polypeptides and polynucleotides, useful for preventing and treating
 PT a disorder associated with increased or decreased expression of RNA
 PT associated proteins.
 XX
 PS Claim 1; Page 103-104; 131pp; English.
 XX
 CC The present sequence represents a human RNA-associated protein. The
 CC expression of RNA-associated proteins is closely associated with
 CC reproductive tissues, nervous tissues, cell proliferation including
 CC cancer, inflammation and immune responses, and so they may be used for
 CC diagnosis, treatment or prevention of cell proliferative, diseases and
 CC immune/inflammatory disorders, and reproductive disorders. Diseases and
 CC disorders which may be treated include actinic keratosis,
 CC atherosclerosis, arteriosclerosis, bursitis, cirrhosis, hepatitis, mixed
 CC connective tissue disease, myelofibrosis, paroxysmal nocturnal
 CC hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia
 CC and cancers, and trauma
 XX
 SQ Sequence 351 AA;
 Alignment Scores:
 Pred. No.: 6 81e-210 Length: 351
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 98.65% Indels: 0
 DB: 3 Gaps: 0
 US-10-698-228-12 (1-1053) x AAY84439 (1-351)
 QY 1 ATGGGCGACCCGGAAGCGCGGCGGTGGATCAGGATCAGAGATCATCTTCA 60
 Db 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
 QY 61 GACACCAAGAGTGAATGAAGTCAATGAAGAGCCACTCTAGAAAGAGTCTCGC 120
 Db 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
 QY 121 CGSTTTGTCTCTTCCATCCAGTACCCTGATTTGGAAATGTATAAACAGGCACAG 180
 Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetIleGlnAlaGln 60
 QY 181 GCTTCCTTCTGGACAGACAGAGGTTGACTTATCAAGGATCTCCCTCACTGGAACAAG 240
 Db 61 AlaSerPheTrpThrAlaGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80

QY 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTACATCTTAGCCTTTTTCAGCCAGTGTAT 300
 Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
 QY 301 GGAATTGTAAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGAGGTGCAGGTTCAGAGCT 360
 Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCATGAGATGTTCACTCAGAGATCTACAGTTTG 420
 Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 QY 421 CTGATAGACACTTACATCAGATGCCAAGAAAGGCAATTTTATTATTAATGCAATTGAA 480
 Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATGGGCCCTTGGATGGATGAGATAGAGATAAAA 540
 Db 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
 QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTCTGTAGAGAGTCTTCTCTCAGGA 600
 Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGluValPhePheSerGly 200
 QY 601 TCTTTTGTCTATATCTGCTAAAGACAGAGAGTCTTATGCCAGGACTCCTCTTTTCC 660
 Db 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 QY 661 AATGAACTCATCAGCAGATGAAGACTTCTCAGTGTGACTTTGCTGCTGCTGCTGCTGCTCAA 720
 Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 QY 721 TACTTAGTAATAAGCCTTCAGAAAGAGGTGAGGAGATCATTTGATGCTGTCAA 780
 Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
 QY 781 ATTGAGCAGAGTCTTTTAAACAGAGAGCTTCCAGTTCGCTCATTTGGAATCAATTGCATT 840
 Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 QY 841 TTGATGAAACAGTACATGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
 Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 QY 901 AAGTCTTTTCAGCAGAAAATCCTTTTGATTTTATGGAATAACATTTCTTTAGAGGAAAA 960
 Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 QY 961 ACAAAATTTCTTTCAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGGTTTATGCGAGAAACC 1020
 Db 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 QY 1021 ACAGATAACGCTCTTCACTTTGGATGAGATTTT 1053
 Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 RESULT 2
 AAB69050
 ID AAB69050 standard; protein; 351 AA.
 XX
 AC AAB69050;
 XX
 DT 18-APR-2001 (first entry)
 XX
 DE Human ribonucleotide reductase TP53R2H protein sequence SEQ ID NO:1.
 XX
 KW Human; ribonucleotide reductase; cancer; DNA repair; p53.
 OS Homo sapiens.
 XX
 PN WO200100799-A1.
 XX

PD 04-JAN-2001.
 XX PF 27-JUN-2000; 2000WO-JP004189.
 XX PR 28-JUN-1999; 99JP-00181131.
 XX PR 06-JUL-1999; 99JP-00192391.
 XX PR 21-JAN-2000; 2000JP-00017770.
 XX (TAKE) TAKEDA CHEM IND LTD.
 PA (NAKA) NAKAMURA Y.
 XX PI Nakamura Y, Arakawa H, Tanaka H;
 XX N-PSDB; AAF32438.
 DR WPI; 2001-112446/12.
 DR N-PSDB; AAF32438.
 XX Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
 PT diagnosis, treatment and prevention of cancer.
 XX Claim 1; Fig 1-3; 102pp; Japanese.
 PS The present sequence represents a human ribonucleotide reductase
 CC designated TP53R2H. TP53R2H has cytosolic activity. TP53R2H is part of
 CC the DNA repair mechanism and its activity is induced by p53. It can be
 CC used for the treatment, prevention and diagnosis of a wide range of
 CC cancers
 XX Query Match: 98.65%
 XX DB: 4
 SQ Sequence 351 AA;
 Alignment Scores:
 Pred. No.: 6.81e-210 Length: 351
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 98.65% Indels: 0
 DB: 4 Gaps: 0
 US-10-698-228-12 (1-1053) x AAB69050 (1-351)
 QY 1 ATGGGCGACCCGGAAGAGCGCGGCTGGATCAGATGAGAGATCATCTTCA 60
 DB 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSer 20
 QY 61 GACACCAACGAAGTCAATTAAGTCAATTAAGTCAATTAAGTCAATTAAGTCAAT 120
 DB 21 AspThrAsnGluSerGluLeuLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
 QY 121 CGTTTGTTCATCTTCCCAATCCAGTACCTCTCATCTTCAATTAAGTCAATTAAGTCAAT 180
 DB 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrPheMetTyrLysGlnAlaGln 60
 QY 181 GCTTCCTCTGACGACGACGAGAGGTGACTATCAAGATCTCCCTCACTGGAACAAG 240
 DB 61 AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
 QY 241 CTTAAGCAGATGAGAGTACTTCTCTCATCTTCAATTAAGTCAATTAAGTCAATTAAGTCAAT 300
 DB 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
 QY 301 GGAATTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCAGAGGCT 360
 DB 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 QY 361 CGCTGTTTCTATGGCTTCAAAATTTCTATCGAATGTTTCACTCAGAGATGTACAGTTTG 420
 DB 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 QY 421 CTGATGACACTTACATCAGATCCCAAGAAAGGAATTTTATTTATGCAATTGAA 480
 DB 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 QY 481 ACCATGCCCTTGTAAAGAAAAAGCAGATTGGCTTGGATGGATAGCAGATGAAAAA 540

Db 161 ThrMetProTyrValLysLysLysAlaAspTrrAlaLeuArgTrrPilleAlaAspArgLys 180
 QY 541 TCTACTTTTGGGAAAGAGTGGCTTTGCTGTAGAGGAGTCTTCTTCTCAGGA 600
 Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
 QY 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACATTTTCC 660
 Db 201 SerPheAlaAlaIlePheTrrPheLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 QY 661 AATGAATCATCAGCAGAGATGAAGACTTCTACGTGACTTTGCTTGGCTGATGTTCAA 720
 Db 221 AsnGluLeuIleSerArgAspGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 QY 721 TACTTAGTAATAGCTTCCAGAGAAAGGTGAGGAGATCATTTGTTGCTGTCTCAA 780
 Db 241 TyrLeuValAsnLysProSerGluArgValArgGluIleIleValAspAlaValLys 260
 QY 781 ATTGAGCAGGAGTTTAAACAGAGCTTCCAGTTCCTCATTTGGAATGAATTCATT 840
 Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 QY 841 TTGATGAAACAGTACATTTAGTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 900
 Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuGlyPheSer 300
 QY 901 AAGTTTTCAGGACAGAAATCCTTTGATTTATGGAACACATTTCTTTAGAGGAAA 960
 Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 QY 961 ACAATTTCTTTCAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTATGCGCAGAAACC 1020
 Db 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 QY 1021 ACAGATAACGTCTTTCACCTTGGATGAGATTTT 1053
 Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 RESULT 3
 AAB93497
 ID AAB93497 standard; protein; 351 AA.
 XX AAB93497;
 AC AAB93497;
 DT 26-JUN-2001 (first entry)
 DE Human protein sequence SEQ ID NO:12811.
 XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX Homo sapiens.
 PN EP1074617-A2.
 PD 07-FEB-2001.
 XX 28-JUL-2000; 2000BP-00116126.
 PF 29-JUL-1999; 99JP-00248036.
 XX 27-AUG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.
 XX (HELI-) HELIX RES INST.
 XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WPI; 2001-318749/34.
 DR Primer sets for synthesizing polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or

PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.
 PS Claim 8; SEQ ID NO 12811; 2537pp + Sequence Listing; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention
 XX
 SQ Sequence 351 AA;

Alignment Scores:

Pred. No.: 351
 Score: 1817.00 Matches: 350
 Percent Similarity: 100.00% Conservatives: 1
 Best Local Similarity: 99.72% Mismatches: 0
 Query Match: 98.43% Indels: 0
 DB: 4 Gaps: 0

US-10-698-228-12 (1-1053) x AAB93497 (1-351)

QY 1 ATGGCGACCCGGAAGCGCGGCTGGATCAGATGAGATCATCTTCA 60
 DB 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSer 20
 QY 61 GACACCAAGCAAGTCAATGAATGCAATGAGAGCCACTCTAGAAAGAGTTCTCC 120
 DB 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
 QY 121 CGTTTGTCTATCTTCCATCCAGTACCCTGATATTGGAAATGTATAACAGGCACAG 180
 DB 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
 QY 181 GCTTCCTTTTGACACGACAGAGGTTGACTTATCAAGGATCTCCCTCACTGGACCAAG 240
 DB 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
 QY 241 CTTAAGCAGATGAGAGTACTTCTCTCATCTTACATCTTACCCCTTTTTCAGCCAGTAT 300
 DB 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
 QY 301 GGAATTTGTAATGAAATTTGGTGGAGCCCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
 DB 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 QY 361 CGCTGTTTCTATGGCTTTCAAAATTTCTATCGAATGTTTCACTCAGAGATGTACAGTTTG 420
 DB 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 QY 421 CTGATAGACATTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 480
 DB 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 QY 481 ACCATGCCCTATGTAAAGAAAAAGCAGATGGCCCTTCGATGGATAGCAGTAGAAAA 540

DB 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
 QY 541 TCTACTTTGGGCAAGAGAGTGGCTTTCCTGCTCTAGAGGAGTTCCTTCTCAGGA 600
 DB 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
 QY 601 TCTTTTGTCTCTATATTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCAGCTTTTCC 660
 DB 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 QY 661 AATGAATCATCAGCAGAGATGAAGACTTCACGTGACTTTGTGCTGCTGATGTTCCAA 720
 DB 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 QY 721 TACTTAGTAATAGCTTTCAGAGAAGGGTCAGGGAGCATCTGTGCTGCTCAAA 780
 DB 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
 QY 781 ATTGAGCAGGAGTCTTTTAAACAGAGCCCTGCGAGTTCGCTTCATTTGGAATGCAATTGCATT 840
 DB 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyValAsnCysIle 280
 QY 841 TTGATGAACAGTACATTGAGTTTGTAGTGCACAGATTACTTGTGGAACCTTGGATTCTCA 900
 DB 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 QY 901 AAGTTTTCAGCAGCAAAAATCCTTTTGATTTATGAAAACATTTCTTTAGAGAGAAA 960
 DB 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 QY 961 ACAAAATTTCTTCAGAAAACAGTTTTCAGAGTATCAGCGTTTTGTCAGTTATGCGCAGAAACC 1020
 DB 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 QY 1021 ACAGATAACGCTTTCACCTTTGGATGCGAGTTTT 1053
 DB 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 4
 AAB69051
 ID AAB69051 standard; protein; 389 AA.
 XX
 AC AAB69051;
 XX
 DT 18-APR-2001 (first entry)
 XX
 DE Human R2 protein sequence SEQ ID NO:5.
 DE
 KW Human; ribonucleotide reductase; cancer; DNA repair; p53.
 XX
 OS Homo sapiens.
 XX
 PN WO200100799-A1.
 XX
 PD 04-JAN-2001.
 XX
 PF 27-JUN-2000; 2000WO-JP004189.
 XX
 PR 28-JUN-1999; 99JP-00181131.
 PR 06-JUL-1999; 99JP-00192391.
 PR 21-JAN-2000; 2000JP-00017770.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 PA (NAKA/) NAKAMURA Y.
 XX
 PI Nakamura Y, Arakawa H, Tanaka H;
 XX
 DR WPI; 2001-112446/12.
 XX
 PT Ribonucleotide reductase involved in DNA repair and DNA encoding it, for
 PT diagnosis, treatment and prevention of cancer.
 XX

conditions especially relating to prostate cancer. (I) and its expression products are used in the diagnostic test to assay for presence of cancer e.g. in tissue sections, in biopsy sample, in total RNA, in lymph, in blood etc. (II) is useful for assessing cancer e.g. to determine the type of cancer, its stage of development, the nature of genetic defect, etc. The polypeptide encoded by (I) can be used as target for therapy or drug discovery. (I) can also be used for expressing the polypeptide and thus for searching specific binding partners of the polypeptide. (I) is useful in therapeutic applications to treat prostate cancer. The identification of specific genes, and groups of genes, expressed in pathways physiologically relevant to prostate cancer permits the definition of functional and disease pathways and the delineation of targets in these pathways which are useful in diagnostic, therapeutic, and clinical applications. This is the amino acid sequence of a protein differentially regulated in prostate cancer

Db	289	ValArgGluLeuIleIleAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu	308
Qy	811	CCAGTTGGGCTCATTCGAATGAATTCATTTTCATGAAAAACAGTACATTCAGTTTGTAGCT	870
Db	309	ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrlleGluPheValAla	328
Qy	871	GACAGATTACTTGTGGAACTTGATTCCTCAAAAGGTTTTTCAGGCAGAAAAATCCTTTTGAT	930
Db	329	AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp	348
Qy	931	TTTTATCGAAAAACATTTCTTTAGNAGGAAAAACAAATTTCTTTGAGAAACGAGTTTCAGAG	990
Db	349	PheMetGluAsnIleSerLeuGluGlyLysThrAsnPheGluLysArgValGlyGlu	368
Qy	991	TATCAGCGTTTTGCAGTTATGGCAGAAACACACAGATAACGTCCTTCACCTTGGATGCAGAT	1050
Db	369	TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp	388
Qy	1051	TTT 1053	
Db	389	Phe 389	
RESULT 6			
ADJ68753			
ID	ADJ68753	standard; protein; 389 AA.	
XX	ADJ68753;		
XX	06-MAY-2004	(first entry)	
XX	Human heat mitochondrial protein as a therapeutic target	SeqID559.	
XX	mitochondrial; human; screening assay; diabetes mellitus;		
KW	Huntington's disease; osteoarthritis;		
KW	leber's hereditary optic neuropathy; LHON;		
KW	mitochondrial encephalopathy lactic acidosis and stroke; MELAS;		
KW	myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;		
KW	neuroprotective; nontropic; antidiabetic; anticonvulsant; antiarthritic;		
XX	osteopathic; ophthalmological; cytostatic.		
XX	Homo sapiens.		
XX	WO2003087768-A2.		
XX	23-OCT-2003.		
XX	04-APR-2003; 2003WO-US010870.		
XX	12-APR-2002; 2002US-0372843P.		
PR	17-JUN-2002; 2002US-0389987P.		
PR	20-SEP-2002; 2002US-0412418P.		
XX	(MITO-) MITOKOR.		
PA	(BUCK-) BUCK INST AGE RES.		
PA	Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;		
PI	Warnock DE;		
XX	WPI; 2003-845369/78.		
DR			
XX	Identifying a mitochondrial target for drug screening assays and for		
PT	treating diseases associated with altered mitochondrial function,		
PT	comprises detecting a modified polypeptide in a sample and correlating		
PT	with the disease.		
XX	Claim 1; SEQ ID NO 559; 180pp; English.		
XX	This invention relates to novel mitochondrial targets that can be used		
CC	for therapeutic intervention in treating a disease associated with		
CC	altered mitochondrial function. Specifically, it refers to a method for		
CC	identifying proteins of the human heart mitochondrial proteome that are		
CC	useful for drug screening assays, as well as therapeutic targets. The		

CC present invention describes a method for identifying such proteins that
CC can be used in the treatment of various diseases associated with altered
CC mitochondrial function including diabetes mellitus, Huntington's disease,
CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
CC encephalopathy, lactic acidosis and stroke (MELAS), myoclonic epilepsy
CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
CC compositions have neuroprotective, neurotropic, antidiabetic,
CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and
CC cytostatic activities. This polypeptide sequence is a human heart
CC mitochondrial protein of the invention.

XX SQ Sequence 389 AA;

Alignment Scores:
Pred. No.: 1,42e-165 Length: 389
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: 7 Gaps: 0

US-10-698-228-12 (1-1053) x ADJ68753 (1-389)

QY	91	GAAGAGCCACTCTTAAGAAAGAGTTCTCGCGGTTTGTTCATCTTTCCATCCAGTACCCT	150
DB	69	AspGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGluTyrHis	88
QY	151	GATATTGGAATATGATAACAGGACAGGCTTCTTCGGACACAGAGGTTGAC	210
DB	89	AspIleTrpGlnMetTyrIleValAlaGluAlaSerPheTrpThrAlaGluGluValAsp	108
QY	211	TTATCAAGGATCTCCCTCACTGGACAGCTTAAGACAGATGAGAGTACTCTCT	270
DB	109	LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer	128
QY	271	CACATCTTAGCCCTTTTTCAGCCAGTGTGAATTTGTAATGAAATTTGGTGGAGCGC	330
DB	129	HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg	148
QY	331	TTTAGTCAGAGGTGAGGTTCAGAGGCTCGCTGTTTATGGCTTTCAAAATTCATC	390
DB	149	PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet	168
QY	391	GAGATGTTCACTCAGAGATGTACAGTTTTCGTATAGACACTTACATCAGAGATCCCAAG	450
DB	169	GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys	188
QY	451	AAAAGGGAATTTTATTATGCAATTTGAAACCATCCCTATGTTAAGAAAAAGCAGAT	510
DB	189	GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp	208
QY	511	TGGGCTTCGATGAGATAGACATAGAAATCTACTTTTGGGAAAGAGTGGGCTTT	570
DB	209	TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaPhe	228
QY	571	GCTGCTGATAGAGGATTTCTTCTCAGCATCTTTGCTGCTATATTCGGCTAAGAAG	630
DB	229	AlaAlaValGluGlyIlePheSerGlySerPheAlaSerIlePheTrpLeuLysLys	248
QY	631	AGAGGTCTTATCCAGAGCTCACTTTTCCATGAACTCATCAGCAGATGAAGACTT	690
DB	249	ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu	268
QY	691	CAGTGTGACTTGTCTGCTGATGTTTCCAAATCTAGTAAATAGCCTTTCAGAAGAAG	750
DB	269	HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg	288
QY	751	GTCAGGAGATCATTTGATGCTGTCAAAATAGCAGAGGTTTATACAGAGCCTTG	810
DB	289	ValArgGluIleIleIleAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu	308
QY	811	CCAGTTGGCTCATTTGGAATGAATTCATTTTGTGTAAGACAGTACATTCAGTTGTAGCT	870

DB	309	ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla	328
QY	871	GACAGATTACTTGTGGAACTTGGATTCTCAAGGTTTTTTCAGGCAGAAATCCTTTTGAT	930
DB	329	AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp	348
QY	931	TTTATGAAAAACATTTCTTTAGAGGAAAAACAAATTTCTTTGAGAAACGAGTTTCAGAG	990
DB	349	PheMetGluAsnIleSerLeuGluGlyThrAsnPhePheGluLysArgValGlyGlu	368
QY	991	TATCAGCGTTTTCAGTTATGGCGAGAAACACAGATACGTCCTTCACCTTGGATGCGAT	1050
DB	369	TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp	388
QY	1051	TTT 1053	
DB	389	Phe 389	
RESULT 7			
ADJ66564			
ID	ADJ66564	standard; protein; 389 AA.	
XX	ADJ66564;		
XX	06-MAY-2004	(first entry)	
XX	Ribonucleoside-diphosphate reductase M2 chain for anti-cancer complex.		
XX	neuroprotective; cytostatic; gene therapy; protein complex;		
KW	cellular network; cancer; neurodegenerative disease; drug target.		
XX	Homo sapiens.		
XX	WO2004009622-A2.		
XX	29-JAN-2004.		
XX	18-JUL-2003; 2003WO-EP007835.		
XX	19-JUL-2002; 2002EP-00016109.		
PR	19-JUL-2002; 2002EP-00016111.		
PR	19-JUL-2002; 2002EP-00016123.		
PR	19-JUL-2002; 2002EP-00016128.		
PR	22-JUL-2002; 2002EP-00016427.		
XX	(CELL-) CELLZOME AG.		
XX	Merino A, Boummeester T, Bauer A, Drewes G, Marzioch M, Kruse U;		
PI	Superti-Furga G, Eberhard D, Ruffner H, Hobson S, Helftenbein G;		
PI	Cruciat C;		
XX	WPI; 2004-123372/12.		
DR	New protein complexes of cellular networks underlying the development of		
XX	cancer and other diseases, useful for diagnosing and/or treating		
PT	neurodegenerative diseases or cancer, and in drug screening.		
PT	Disclosure; SEQ ID NO 94; 809pp; English.		
XX	The invention relates to a protein complex of cellular networks		
CC	underlying the development of cancer and other diseases. The complex (I)		
CC	comprises at least one first and second proteins selected from any of the		
CC	proteins listed in the specification, or their functionally active		
CC	derivatives, fragments, homologues or variants, the variants being		
CC	encoded by a nucleic acid that hybridizes to the nucleic acid encoding		
CC	the protein under low stringency conditions. A complex (II) comprises at		
CC	least two of the second proteins, where the low stringency conditions		
CC	comprise hybridization in a buffer comprising 35% formamide, 5 x SSC, 50		
CC	mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.02% PVP, 0.02% BSA, 100 microg/ml		
CC	denatured salmon sperm DNA, and 10% (wt/vol) dextran sulfate for 18-20		
CC	hours at 40 deg C, washing in a buffer consisting of 2 x SSC, 25 mM Tris-		
CC	HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS for 1.5 hours at 55 deg C, and		
CC	washing in a buffer consisting of 2 x SSC, 25 mM Tris-HCl (pH 7.4), 5 mM		

CC EDTA, and 0.1% SDS for 1.5 hours at 60 deg C. The composition and methods
CC are useful in diagnosing or treating diseases and disorders, preferably
CC neurodegenerative diseases. These may also be used as a drug target or in
CC manufacturing a medicament for the treatment or prevention of the above-
CC mentioned diseases or disorders. The composition may also be used for
CC treating cancer. This sequence represents one of the proteins of the
CC complex of the invention.

XX SQ Sequence 389 AA;

Alignment Scores:
Pred. No.: 1.42e-165 Length: 389
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: 8 Gaps: 0

US-10-698-228-12 (1-1053) x ADJ66564 (1-389)

QY 91 GAAGAGCCACTCTAAGAAGAGTTCTCGCGGTTTGTTCATCTTCCCAATCCAGTACCCT 150
DB 69 ASPGLUProLeuLeuArgGluAsnProArgArgPheValIlePheProIleGluTyrHis 88
QY 151 GATATTGGAAATGTATAACAGGACAGAGGCTTCTCTGGACAGACAGAGGTTGAC 210
DB 89 AsplleTPrGlnMetTyrIleLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 108
QY 211 TTATCAAGAGTCTCCCTCACTGGAAACAAGCTTAAAGCAGAGTACTTCATCTCT 270
DB 109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128
QY 271 CACATCTTAGCTTTTTCGACGCGCTGATGGAATTAATGAAATTTGGTGGAGCGC 330
DB 129 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148
QY 331 TTATGTCAGAGTGTGAGTTCAGAGGCTCGCTTTCTATGGCTTTCAAATTTCTCATC 390
DB 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
QY 391 GAGATGTTCTACGAGATGTACAGTTTCTGATGACACTTACATCAGAGATCCCAAG 450
DB 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
QY 451 AAAAGGGAATTTTATTAATGCAATTAAGAACCATGCCCTATGTTAAAGAAAAGCAGAT 510
DB 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValIleLysLysAlaAsp 208
QY 511 TGGGCTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 570
DB 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaAlaPhe 228
QY 571 GCTGCTGTAGAAGAGTCTTCTCTCAGGATCTTTTCTGCTATATTTGGCTTAAGAAG 630
DB 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 248
QY 631 AGAGTCTTATGCGAGGACTACTTTTTCATGAACTCATCAGCAGAGATGAAGACTT 690
DB 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
QY 691 CACTGTGACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 750
DB 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
QY 751 GTGAGGAGATCATTTGATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 810
DB 289 ValArgGluIleIleIleAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 308
QY 811 CCAGTTGGGCTCATTTGGAATGAATTCATTTTGAATGAAACAGTACATTTGATTTGTAGCT 870
DB 309 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 328
QY 871 GACAGATTACTTGTGAACTTGATTTCTCAAGGTTTTCAGGACAGAAATTCCTTTTGAT 930

DB 329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348
QY 931 TTTATCGAAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACGAGTTTCAGAG 990
DB 349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPheGluLysArgValGlyGlu 368
QY 991 TATCAGCGTTTTCAGTTATGCGGAAACCCACAGATAACCTTTCACCTTGGATGCCAGAT 1050
DB 369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388

QY 1051 TTT 1053
DB 389 Phe 389
RESULT 8
ADK70435
ID ADK70435 standard; protein; 389 AA.
XX ADK70435;
XX 06-MAY-2004 (first entry)
XX Respiratory disease differentially expressed protein #1.
XX cytotstatic; respiratory; antiasthmatic; Gene therapy;
XX differential gene expression; respiratory disorder; lung cancer;
XX chronic obstructive pulmonary disease; emphysema; asthma.
XX Homo sapiens.
XX WO2003101283-A2.
XX 11-DEC-2003.
XX 02-JUN-2003; 2003WO-US017409.
XX 04-JUN-2002; 2002US-0386005P.
XX (INCY-) INCYTE CORP.
XX Rickert PK, Krasnow R;
XX WPI; 2004-042945/04.
XX New combination comprising cDNAs and proteins that are differentially
XX expressed in respiratory disorders, useful for diagnosing or treating
XX respiratory diseases e.g. lung cancer, chronic obstructive pulmonary
XX diseases or asthma.
XX Claim 14; SEQ ID NO 171; 343pp; English.

CC The invention relates to cDNA sequences that are differentially expressed
CC in respiratory disorders or their complements or encoded proteins. The
CC cDNAs and proteins are useful for diagnosing, treating or monitoring
CC treatment of a subject with a respiratory disease including lung cancer,
CC chronic obstructive pulmonary diseases, emphysema or asthma. The protein
CC is also useful for screening molecules or compounds to identify at least
CC one ligand which specifically binds the protein. It is also useful for
CC preparing and purifying a polyclonal or monoclonal antibody. This
CC sequence corresponds to a protein of the invention.

XX SQ Sequence 389 AA;
Alignment Scores:
Pred. No.: 1.42e-165 Length: 389
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: 8 Gaps: 0

US-10-698-228-12 (1-1053) x ADK70435 (1-389)

```
QY 91 GAAGGCCACTCTAAGAAAGAGTTCTCGCGGTTTGTTCATCTTTCCAAATCCAGTACCCT 150
   ::::::::::::::::::::
Db 69 AspGluProLeuLeuArgGluAsnProArgArgPheValIlePheProIleGluTyrHis 88
QY 151 GATATTGGAAAATGATATAACAGGCACAGGCTTCTTCGACACAGAGAGGTTGAC 210
   ::::::::::::::::::::
Db 89 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 108
QY 211 TTATCAAGGATCTCCCTCACTCGAACAAGCTTAAGACAGATCAGAAGTACTCTCT 270
   ::::::::::::::::::::
Db 109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128
QY 271 CACATCTAGCCTTTTTCAGCCAGTGTAAATGTAATGAAAATTTGGTGGAGCGC 330
   ::::::::::::::::::::
Db 129 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148
QY 331 TTATGTCAGAGGTGCAGGTTCCAGAGGCTCGCTTCTATGGCTTCAAAATTCATC 390
   ::::::::::::::::::::
Db 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
QY 391 GAGATGTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAGATCCCAAG 450
   ::::::::::::::::::::
Db 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
QY 451 AAAAGGGAATTTTATTATGCAATTTGAAACCATCGCTATCTTAAAGAAAAGCAGAT 510
   ::::::::::::::::::::
Db 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 208
QY 511 TGGGCTTCGCGATGATAGATAGATAGAAAATCTACTTTTGGGAAAAGAGTGGGCTTT 570
   ::::::::::::::::::::
Db 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValValAlaPhe 228
QY 571 GCTGCTGTAGAGGATTTTCTTCAGGATCTTTTGTCTGCTATATCTGGCTAAAGAG 630
   ::::::::::::::::::::
Db 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 248
QY 631 AGAGGCTTATGCCAGGATCTACTTTTCCNATGAATCTATCAGCAGATGAGACATT 690
   ::::::::::::::::::::
Db 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
QY 691 CACTGTGACTTTGCTGCTGATGTTCCTCAATCTAGTAAATTAAGCTTCAGAAAGAGG 750
   ::::::::::::::::::::
Db 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
QY 751 GTCAGGAGATCATTTGATGCTGTCAAAATTCAGCAGAGTTTAAACAGAGCCTTG 810
   ::::::::::::::::::::
Db 289 ValArgGluIleIleAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 308
QY 811 CCAGTTGGGCTCATTTGGAATGAATTCATTTTGAATGAACAGTACATTTGATGCT 870
   ::::::::::::::::::::
Db 309 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 328
QY 871 GACAGATTACTTGTGAATCTGATTTCTCAAGGTTTTCAGCGCAAAATCTTTTGAT 930
   ::::::::::::::::::::
Db 329 AspArgLeuMetCeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348
QY 931 TTATGGAAAACATTTCTTTAGAGGAAAACAAATTTCTTTGAGAAAACGAGTTTCAGAG 990
   ::::::::::::::::::::
Db 349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368
QY 991 TATCAGGTTTGGAGTTATGGCAGAAACCAACAGATAAGCTTTCATCCTTGGATGAGAT 1050
   ::::::::::::::::::::
Db 369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
QY 1051 TTT 1053
   |||
Db 389 Phe 389
RESULT 9
ADN03789
ID ADN03789 standard; protein; 389 AA.
XX
```

```
AC ADN03789;
XX 01-JUL-2004 (first entry)
XX Antipsoriatic protein sequence #91.
XX antipsoriatic; gene therapy; psoriasis; diagnosis.
XX Homo sapiens.
XX WO2004028479-A2.
XX 08-APR-2004.
XX 25-SEP-2003; 2003WO-US030907.
XX 25-SEP-2002; 2002US-0414006P.
XX (GETH ) GENENTECH INC.
XX Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
XX Wu TD;
XX WPI; 2004-305105/28.
XX N-PSDB; ADN03788.
XX New PRO nucleic acid or polypeptide, useful for preparing a
XX pharmaceutical composition for diagnosing or treating psoriasis in a
XX mammal.
XX Claim 9; SEQ ID NO 183; 3069pp; English.
XX The invention relates to novel polynucleotide and polypeptides for
XX treating psoriasis or a sequence having at least 80% identity to the
XX above sequences. The nucleic acid is useful for preparing a composition
XX for diagnosing or treating psoriasis in a mammal. This sequence
XX corresponds to one of the polypeptides of the invention.
XX Sequence 389 AA;
SQ Alignment Scores:
Pred. No.: 1,42e-165 Length: 389
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: Gaps: 8
US-10-698-228-12 (1-1053) x ADN03789 (1-389)
QY 91 GAAGGCCACTCTAAGAAAGAGTTCTCGCGGTTTGTTCATCTTTCCAAATCCAGTACCCT 150
   ::::::::::::::::::::
Db 69 AspGluProLeuLeuArgGluAsnProArgArgPheValIlePheProIleGluTyrHis 88
QY 151 GATATTGGAAAATGATATAACAGGCACAGGCTTCTTCGACACAGAGAGGTTGAC 210
   ::::::::::::::::::::
Db 89 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 108
QY 211 TTATCAAGGATCTCCCTCACTCGAACAAGCTTAAGACAGATCAGAAGTACTCTCTCT 270
   ::::::::::::::::::::
Db 109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128
QY 271 CACATCTAGCCTTTTTCAGCCAGTGTAAATGTAATGAAAATTTGGTGGAGCGC 330
   ::::::::::::::::::::
Db 129 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148
QY 331 TTATGTCAGAGGTGCAGGTTCCAGAGGCTCGCTTCTATGGCTTCAAAATTCATC 390
   ::::::::::::::::::::
Db 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
QY 391 GAGATGTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAGATCCCAAG 450
   ::::::::::::::::::::
Db 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
QY 451 AAAAGGGAATTTTATTATGCAATTTGAAACCATCGCTATCTTAAAGAAAAGCAGAT 510
   ::::::::::::::::::::
Db 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 208
QY 511 TGGGCTTCGCGATGATAGATAGATAGAAAATCTACTTTTGGGAAAAGAGTGGGCTTT 570
   ::::::::::::::::::::
Db 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValValAlaPhe 228
QY 571 GCTGCTGTAGAGGATTTTCTTCAGGATCTTTTGTCTGCTATATCTGGCTAAAGAG 630
   ::::::::::::::::::::
Db 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 248
QY 631 AGAGGCTTATGCCAGGATCTACTTTTCCNATGAATCTATCAGCAGATGAGACATT 690
   ::::::::::::::::::::
Db 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
QY 691 CACTGTGACTTTGCTGCTGATGTTCCTCAATCTAGTAAATTAAGCTTCAGAAAGAGG 750
   ::::::::::::::::::::
Db 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
QY 751 GTCAGGAGATCATTTGATGCTGTCAAAATTCAGCAGAGTTTAAACAGAGCCTTG 810
   ::::::::::::::::::::
Db 289 ValArgGluIleIleAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 308
QY 811 CCAGTTGGGCTCATTTGGAATGAATTCATTTTGAATGAACAGTACATTTGATGCT 870
   ::::::::::::::::::::
Db 309 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 328
QY 871 GACAGATTACTTGTGAATCTGATTTCTCAAGGTTTTCAGCGCAAAATCTTTTGAT 930
   ::::::::::::::::::::
Db 329 AspArgLeuMetCeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348
QY 931 TTATGGAAAACATTTCTTTAGAGGAAAACAAATTTCTTTGAGAAAACGAGTTTCAGAG 990
   ::::::::::::::::::::
Db 349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368
QY 991 TATCAGGTTTGGAGTTATGGCAGAAACCAACAGATAAGCTTTCATCCTTGGATGAGAT 1050
   ::::::::::::::::::::
Db 369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
QY 1051 TTT 1053
   |||
Db 389 Phe 389
RESULT 9
ADN03789
ID ADN03789 standard; protein; 389 AA.
XX
```

```
QY 451 AAAAGGGAATTTTATTATTAATGCAATGAAACCATGCCCTATGTTAGAAAAAGCAGAT 510
DB 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAasp 208
QY 511 TGGGCTTCGATGGATAGCAGATAGAAAATCTACTTTTGGGAAAGAGTGGGCTTT 570
DB 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaPhe 228
QY 571 GCTGCTGTAGAGGATTTCTCTCAGGATCTTTGCTGCTATATCTGGCTAAAGAG 630
DB 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLys 248
QY 631 AGAGGTCTTATGCAGGACTCACTTTTCCAATGAACTCATCAGCAGAGATGAAGACTT 690
DB 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgaspGluGlyLeu 268
QY 691 CACTGTGACTTTGCTGCTGATGTTCCAACTTAGTAATAAGCCTTCAGAAAGAG 750
DB 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
QY 751 CTCAGGAGATCATTTGATGCTGCTCAAAATGACAGAGTTTTTAACAGAGCCTTG 810
DB 289 ValArgGluIleIleLeuAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 308
QY 811 CCAGTTGGCTCATTCGAATGAATTCGATTTGATGAAACAGTACATTCAGTTGTAGCT 870
DB 309 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTrpIleGluPheValAla 328
QY 871 GACAGATTACTTGTGGAATCTGGAATCTCAAAAGTTTTCAGGCAGAAAATCCTTTGAT 930
DB 329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAasp 348
QY 931 TTTATGGAACATTTCTTAGAGAAAACAAATTTCTTCAGAAAACGAGTTTCAGAG 990
DB 349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368
QY 991 TATCAGCGTTTTCAGTTATGGCAGAAACACAGATAACGCTTTCACCTTGGATGCAGAT 1050
DB 369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAaspAlaAasp 388
QY 1051 TTT 1053
DB 389 Phe 389
RESULT 10
ADN04444
ID ADN04444 standard; protein; 389 AA.
XX AC ADN04444;
XX DT 01-JUL-2004 (first entry)
XX DE Antipsoriatic protein sequence #415.
XX KW antipsoriatic; gene therapy; psoriasis; diagnosis.
XX OS Homo sapiens.
XX PN W02004028479-A2.
XX PD 08-APR-2004.
XX PF 25-SEP-2003; 2003WO-US030907.
XX PR 25-SEP-2002; 2002US-0414006P.
XX PA (GETH ) GENENTECH INC.
XX PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
XX WU Wu TD;
XX WPI; 2004-305105/28.
```

```
DR N-PSDB; ADN04443.
XX New PRO nucleic acid or polypeptide, useful for preparing a
PT pharmaceutical composition for diagnosing or treating psoriasis in a
PT mammal.
XX Claim 9; SEQ ID NO 838; 3069pp; English.
XX The invention relates to novel polynucleotide and polypeptides for
CC treating psoriasis or a sequence having at least 80% identity to the
CC above sequences. The nucleic acid is useful for preparing a composition
CC for diagnosing or treating psoriasis in a mammal. This sequence
CC corresponds to one of the polypeptides of the invention.
XX Sequence 389 AA;
SQ
Alignment Scores:
Pred. No.: 1,42e-165 Length: 389
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: 8 Gaps: 0
US-10-698-228-12 (1-1053) x ADN04444 (1-389)
QY 91 GAAGGCCACTCTCTAAGAAAGAGTTCTCGCGGTTGTCATCTTCCATCCAGTACCCT 150
DB 69 AspGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGluThrHis 88
QY 151 GATATTGGAAAAATGTATAACAGGACAGCGTTCCTTCTGGACAGCAGAGAGGTTGAC 210
DB 89 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAasp 108
QY 211 TTATCAAGAGATCCCTCTCATCGAACACAGCTTAACAGCAGATGAGAGTACTTCATCTCT 270
DB 109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128
QY 271 CACATCTTAGCTTTTTCAGCGCAGTGATGGAATGTAATGAAATTTGCTGGAGCGC 330
DB 129 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148
QY 331 TTTAGTCAGGAGGTGCGAGGTTCAGAGGCTCGCTGTTTCTATGGCTTTCAAATTTCTCATC 390
DB 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
QY 391 GAGAAATGTTCACTCAGAGATGTACAGTTTCTGTAGACACTTATCATCAGAGATCCCAAG 450
DB 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
QY 451 AAAAGGGAATTTTATTATTAATGCAATGAAACCATGCCCTATGTTAGAAAAAGCAGAT 510
DB 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAasp 208
QY 511 TGGGCTTCGATGGATAGCAGATAGAAAATCTACTTTTGGGAAAGAGTGGGCTTT 570
DB 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaPhe 228
QY 571 GCTGCTGTAGAGGATTTTCTCTCAGGATCTTTTCTGCTATATCTGGCTAAAGAG 630
DB 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLys 248
QY 631 AGAGGTCTTATGCAGGACTCACTTTTCCAATGAACTCATCAGCAGAGATGAAGACTT 690
DB 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgaspGluGlyLeu 268
QY 691 CACTGTGACTTTGCTGCTGATGTTCCAACTTAGTAATAAGCCTTCAGAAAGAG 750
DB 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
QY 751 GTCAGGAGATCATTTGATGCTGCTCAAAATGACAGAGTTTTTAACAGAGCCTTG 810
DB 289 ValArgGluIleIleLeuAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 308
```

Claim 9; SEQ ID NO 838; 3069pp; English.

The invention relates to novel polynucleotide and polypeptides for treating psoriasis or a sequence having at least 80% identity to the above sequences. The nucleic acid is useful for preparing a composition for diagnosing or treating psoriasis in a mammal. This sequence corresponds to one of the polypeptides of the invention.

QY 931 TTTATGGAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACGAGTTTCAGAG 990
 Db PheMetGluAsnIleSerLeuGluGlyThrAsnPheGluLysArgValGlyGlu 368
 QY 991 TATCAGCGTTTGGAGTTATGTCAGAAACACAGATAAGCTTTCACCTGGATGAGAT 1050
 Db TyrGlnArgMetGlyValMetSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
 QY 1051 TTT 1053
 Db 389 Phe 389
 RESULT 12
 ID ADQ09272 standard; protein; 389 AA.
 XX ADQ09272;
 AC ADQ09272;
 DT 23-SEP-2004 (first entry)
 XX Human RRM2 protein SEQ ID NO:457.
 XX thanatos-associated protein; THAP; THAP responsive gene; THAP family;
 KW THAP responsive element; angiogenesis; inflammation; metastasis; cancer;
 KW apoptosis; cardiovascular disease; neurodegenerative disease; chemokine;
 KW antiangiogenic; antiinflammatory; cardiovascular; cytostatic;
 KW neuroprotective; osteopathic; THAP modulator; THAP synthesis modulator;
 KW human.
 XX Homo sapiens.
 OS
 PN WO2004055050-A2.
 XX
 PD 01-JUL-2004.
 XX
 XX 10-DEC-2003; 2003WO-IB006434.
 PF
 PR 10-DEC-2002; 2002US-0432699P.
 PR 03-JUL-2003; 2003US-0485027P.
 XX
 PA (ENDO-) ENDOCUBE SAS.
 PA (CNRS) CNRS CENT NAT RECH SCI.
 XX
 PI Girard J, Amalric F, Roussigne M, Clouaire T;
 XX
 DR WPI; 2004-525034/50.
 DR N-PSDB; ADQ09273.
 XX
 XX Modulating expression of a Thanatos (death)-Associated Protein (THAP)
 PT responsive gene for preventing or treating e.g. cancer or inflammation,
 PT comprises modulating the interaction of a THAP polypeptide with a nucleic
 PT acid.
 XX
 PS Example 47; SEQ ID NO 457; 612pp; English.
 XX
 CC The present invention describes a method for modulating the expression of
 CC a thanatos (death)-associated protein (THAP) responsive gene. The method
 CC comprises modulating the interaction of a THAP-family polypeptide or its
 CC biological fragment with a nucleic acid, and so enhancing or repressing
 CC the expression of the THAP responsive gene. Also described: (1) a method
 CC of modulating the expression of a gene responsive to a THAP/chemokine
 CC complex; (2) a pharmaceutical composition comprising a THAP responsive
 CC element in a pharmaceutical carrier; (3) a transcription factor decoy
 CC consisting essentially of a THAP responsive element; (4) a cell
 CC comprising a transcription factor decoy described above; (5) methods of
 CC modulating the interaction between a nucleic acid and a THAP-family
 CC polypeptide or its biological fragment, or a nucleic acid and a
 CC THAP/chemokine complex; (6) a vector packaging cell line comprising a
 CC cell comprising a viral vector which comprises a promoter operably linked
 CC to a nucleic acid encoding a THAP-family polypeptide or its biological
 CC fragment; (7) a method of constructing a cell which expresses a
 CC recombinant THAP-family polypeptide; (8) a method of ameliorating
 CC symptoms associated with a condition mediated by a THAP/chemokine complex

CC ; (9) methods of identifying a test compound that modulates transcription
 CC at a THAP responsive element or that modulates the transport of a
 CC chemokine into the nucleus; (10) methods for reducing the symptoms
 CC associated with a condition selected from excessive or insufficient
 CC angiogenesis, inflammation, metastasis of a cancerous tissue, excessive
 CC or insufficient apoptosis, cardiovascular disease and neurodegenerative
 CC diseases; symptoms associated with a condition resulting from the
 CC activity of a chemokine or a THAP-family polypeptide in an individual; or
 CC symptoms associated with transcriptional repression or activation
 CC mediated by a THAP-family polypeptide in an individual; (11) a vector
 CC comprising a THAP responsive promoter operably linked to a nucleic acid
 CC encoding a detectable product; (12) a genetically engineered cell
 CC comprising the vector described above or that expresses a THAP-family
 CC polypeptide or its biological fragment; (13) an in vitro transcription
 CC reaction comprising a nucleic acid comprising a THAP responsive promoter,
 CC ribonucleotides and an RNA polymerase; and (14) an isolated mutant THAP-
 CC family polypeptide that does not bind to a chemokine. The pharmaceutical
 CC composition has antiangiogenic, antiinflammatory, cardiovascular,
 CC cytostatic, neuroprotective and osteopathic activities, and can be used
 CC as a THAP and THAP synthesis modulator. The composition can be used for
 CC modulating the expression of a THAP responsive gene. Modulation is useful
 CC for reducing symptoms of conditions such as excessive or insufficient
 CC angiogenesis, inflammation, metastasis of a cancerous tissue, excessive
 CC or insufficient apoptosis, cardiovascular disease or neurodegenerative
 CC diseases. The present sequence is used in the exemplification of the
 CC present invention.
 XX
 SQ Sequence 389 AA;

Alignment Scores:

Pred. No.:	1,42e-165	Length:	389
Score:	1454.00	Matches:	269
Percent Similarity:	94.08%	Conservative:	33
Best Local Similarity:	83.80%	Mismatches:	19
Query Match:	8	Indels:	0
DB:		Gaps:	0

US-10-698-228-12 (1-1053) x ADQ09272 (1-389)

QY	91	GAAGAGCCACTCTCTAGAAAGAGTTCTCGCGGTTTGTTCATCTTTCCAAATCAGTACCT	150
Db	69	AspGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGluIle	88
QY	151	GATATTGGAAATGTATAACAGGCACAGCTCTCTCGACAGCAGAGAGGTTCCAC	210
Db	89	AspIletrpGlnMetTyrLysLysAlaGluAlaSerPheIleThrAlaGluGluValAsp	108
QY	211	TTATCAAGGATCTCCCTCACTGGACCAAGCTTAAAGCAGATGAGAGTACTTCATCTCT	270
Db	109	LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer	128
QY	271	CACATCTTAGCCCTTTTTCAGCCAGTGATGGAATTTAAATGAAATTTGGTGAGCGC	330
Db	129	HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg	148
QY	331	TTTATGTCAGAGTGTCAGGTTCCAGAGCTCGCTTCTATGGCTTTCAATTCATCTCATC	390
Db	149	PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet	168
QY	391	GAGAAATGTTCCACTCAGAGATGTACAGTTTGCTGATAGACACTTACATCAGAGATCCCAAG	450
Db	169	GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys	188
QY	451	AAAAGGGAATTTTATTTAATGCAATTTGAAACCATGCCCTATGTTAAGAAAAAGCAGAT	510
Db	189	GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysLysAlaAsp	208
QY	511	TGGGCTTGGCATGATAGCAGATAGAAATCTACTTTTGGGGAAAGAGTGGTGCCTTT	570
Db	209	TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaPhe	228
QY	571	GCCTGCTAGAACGAGTTTCTCTCTCAGGATCTTTTGTGCTGCTATATCTCGCTAAAGAG	630

Db 229 AlaAlaValGluGlyLeuPhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 248
QY 631 AGAGGCTTTATGCGCAGGACTCACTTTTCCAAATGAATCATCAGCAGAGATGAAGGACTT 690
Db 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
QY 691 CACTGTGACTTTCCTGCTGATGTTCCAAATACATTAAGTAAATAGCCTTCAGAAAGG 750
Db 269 HisCysAspPheAlaCysLeuMetPheLeuHisLeuValHisLysProSerGluGluArg 288
QY 751 GTCAGGAGATCACTTGTGCTGCTGCTCAAAATGAGCAGGAGTTTTAAACAGAGCCCTG 810
Db 289 ValArgGluIleIleIleAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 308
QY 811 CCAGTTGGCTCATTGGAATGAATGCAATTTTCATGAAACAGTACATTTAGCTTGTAGCT 870
Db 309 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnIryIleGluPheValAla 328
QY 871 GACAGATTACTTGTGGAATGGAATTCCTCAAGGTTTTCAGCGCAGAAATCCTTTTGAT 930
Db 329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348
QY 931 TTTATGGAACAACTTCTTTAGAGGAGAAACAAATTTCTTCAGAAACGAGTTTCAGAG 990
Db 349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368
QY 991 TATCAGCGTTTTCAGTTATGCGCAGAAACACAGATAACTCTTCACCTTGGATGCAGAT 1050
Db 369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
QY 1051 TTT 1053
Db 389 Phe 389

RESULT 13
ABM80258
ID ABM80258 standard; protein; 389 AA.
XX AC ABM80258;
XX DT 18-NOV-2004 (first entry)
XX DE Tumour-associated antigenic target (TAT) polypeptide PRO62893, SEQ:652.
KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW tumour; diagnosis; cell proliferative disorder; breast cancer;
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW central nervous system cancer; bladder cancer; pancreatic cancer;
KW cervical cancer; melanoma; leukaemia; hybridisation probe;
KW chromosome identification; chromosome mapping; gene mapping;
KW gene therapy; cytostatic.
XX OS Homo sapiens.
XX PN WO2004030615-A2.
XX PD 15-APR-2004.
XX PF 29-SEP-2003; 2003WO-US028547.
XX PR 02-OCT-2002; 2002US-0414971P.
XX PA (GETH) GENENTECH INC.
XX PI Wu TD, Zhang Z, Zhou Y;
XX DR WPI: 2004-347921/32.
XX DR N-PSDB; ACN37637.
XX PT New tumor-associated antigenic target polypeptides and nucleic acids,
XX PT useful in preparing a medicament for treating or detecting a
XX PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
XX PT prostate cancer or tumor.

XX

PS Claim 12; SEQ ID NO 652; 7273pp; English.

CC The invention relates to human tumour-associated antigenic target (TAT) polypeptides, and their related nucleic acids. The TAT polypeptides are overexpressed in cancer tissues compared to normal tissues, and may thus serve as effective targets for the diagnosis and treatment of cancer in mammals. The invention also relates to nucleic acid and polypeptide sequences at least 80% identical to the TAT nucleic acids and polypeptides; expression vectors and host cells comprising a TAT nucleic acid; an antibody specific for a TAT polypeptide; a peptide or organic molecule which binds to a TAT polypeptide; fusion proteins comprising a TAT polypeptide; and methods and compositions for the treatment or diagnosis of cancer in mammals. TAT polypeptides, nucleic acids, antibodies, antagonists, binding molecules, and compositions are useful for diagnosing or treating a cell proliferative disorder associated with increased TAT expression, particularly cancers such as breast cancer, colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder cancer, pancreatic cancer, cervical cancer, cancers of the central nervous system, melanoma and leukaemia. TAT nucleic acids may further be used as hybridisation probes, in chromosome and gene mapping, in chromosome identification and in gene therapy. The present sequence represents a TAT polypeptide of the invention

SQ Sequence 389 AA;

Alignment Scores:

Pred. No.:	1.42e-165	Length:	389
Score:	1454.00	Matches:	269
Percent Similarity:	94.08%	Conservative:	33
Best Local Similarity:	83.80%	Mismatches:	19
Query Match:	78.76%	Indels:	0
DB:	8	Gaps:	0

US-10-698-228-12 (1-1053) x ABM80258 (1-389)

QY 91 GAAGAGCCACTCTCTAAGAAAGAGTTCTCGCGGTTTGTTCATCTTCCAAATCCAGTACCCT 150

Db 69 AspGluProLeuLeuArgGluAsnProArgArgPheValIlePheProIleGluIryHis 88

QY 151 GATATTTGGAAAATGTATAAACAGCAGCAGCTTCTCTCGACAGCAGAGAGGTTGAC 210

Db 89 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValasp 108

QY 211 TTATCAAAGATCTCCCTACTGTGAAACAGCTTAAAGCAGATGAGAGTACTTCTCT 270

Db 109 LeuSerLysAspIleGlnHisTrpGluSerLysProGluGluArgTyrPheIleSer 128

QY 271 CACATCTTAGCCTTTTTCAGCCAGTGATGGAATTTGTAATGAAATTTGTTGGAGCGC 330

Db 129 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148

QY 331 TTTAGTCAGGAGGTGCGAGGTTCCAGAGCTCGCTGTCTTCTATGGCTTTCAAATTTCTCATC 390

Db 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168

QY 391 GAGAGTTTCACTCAGAGATGTACAGTTGCTGTATAGACACTTACATCAGAGATCCCAAG 450

Db 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188

QY 451 AAAAGGGAATTTTATTATTAATGCAATTTGAAACCATGCCCTATGTTAAGAAAAACAGAT 510

Db 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 208

QY 511 TGGGCTTTCGATGATAGCAGATAGAAAATCTACTTTTGGGGAAGAGTGGTGCCTTT 570

Db 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValValAlaPhe 228

QY 571 GCTGCTGTAGAAGAGTTTCTTCTCAGGATCTTTTCTGCTATATTTCTGCTGCTAAAGAG 630

Db 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 248

QY 631 AGAGTCTTATGCCAGGACTCACTTTTCCAACTCACTCATCAGCAGAGATGAGGACTT 690

```

Db 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
|||
QY 691 CACTGTGCTTGGCTGCTGATGTTCCAACTAGTAATAAGCCTTCAGAGAAAGG 750
|||
Db 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
|||
QY 751 GTCAGGAGATCATTTGTCATGCTGTCAAAATTTGACGAGAGTTTAAACAGAGCCTTG 810
|||
Db 289 ValArgGluIleIleIleAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 308
|||
QY 811 CCAGTTGGCCTCATTTGAATGAATGATTTTGATGAAACAGPACATTTGAGTTGTAGCT 870
|||
Db 309 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnIleGluPheValAla 328
|||
QY 871 GACAGATTACTTGTGGAATCTGATCTCAAGGTTTTCAGCGACAAATCTTTGAT 930
|||
Db 329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348
|||
QY 931 TTTATGGAACCAATTTCTTAGAGGAAACACAAATTTCTTGAGAAACGAGTTTCAGAG 990
|||
Db 349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368
|||
QY 991 TATCAGCGTTTTCAGTTATGCGAGAAACACAGATPACGTTCTCACCTTGGATGCAGAT 1050
|||
Db 369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
|||
QY 1051 TTT 1053
|||
Db 389 Phe 389
|||
RESULT 14
AAB43902
ID AAB43902 standard; protein; 413 AA.
XX
AC AAB43902;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human cancer associated protein sequence SEQ ID NO:1347.
XX
KW Human; cancer associated gene; cancer antigen; detection; cancer;
KW diagnosis; cytostatic; proliferative; vulnerable; immunomodulator;
KW antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
KW antinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
KW dermatological; neuroprotective; thrombolytic; coagulant; neutropenic;
KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW allergic reaction; graft versus host disease; organ rejection;
KW haemostatic; thrombolytic; cardiovascular disorder; infection;
KW neurological disease; drug screening.
XX
OS Homo sapiens.
XX
PN WO200055350-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US005882.
XX
PR 12-MAR-1999; 99US-0124270P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI: 2000-587533/55.
XX
DR N-PSDB; AAC78111.
XX
PT Novel isolated nucleic acids comprising sequences encoding peptides
PT useful for treating or diagnosing e.g. cancer.
XX
PS Claim 11; Page 2000-2002; 2352pp; English.

```

```

XX
CC AAC77607 to AAC78448 encode the human cancer associated proteins given in
CC AAB43998 to AAB44239. The proteins can have activities based on the
CC tissues and cells the genes are expressed in. Example of activities
CC include: cytostatic; proliferative; vulnerable; immunomodulator;
CC antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
CC antinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
CC neutropenic; vasotropic; antipsoriatic and antiangiogenic. The
CC polynucleotides and polypeptides can be used for preventing, treating or
CC ameliorating medical conditions and diagnosing pathological conditions.
CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC the present invention may be used to treat immune disorders by activating
CC or inhibiting the proliferation, differentiation or mobilisation of
CC immune cells, to treat disorders of haematopoietic cells, autoimmune
CC disorders, allergic reactions, graft versus host disease and organ
CC rejection, modulate haemostatic or thrombolytic activity, modulate
CC inflammation, cancers, cardiovascular disorders, neurological disease and
CC bacterial or viral infections. The peptides, nucleotides, antibodies,
CC agonists and antagonists may be also be used in drug screens. AAC78449 to
CC AAC78457 and AAB44240 represent sequences used in the exemplification of
CC the present invention
XX
SQ Sequence 413 AA;

```

```

Alignment Scores:
Pred. No.: 1,46e-165 Length: 413
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservatives: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: Gaps: 0

```

US-10-698-228-12 (1-1053) x AAB43902 (1-413)

```

QY 91 GAAGAGCCACTCTCTAAGAAAGAGTTCTCGCGGTTTCTCATCTTTTCCAAATCCAGTACCCT 150
:::|||||
Db 93 AspGluProLeuLeuArgGluAsnProArgArgPheValIlePheProIleGluTyrHis 112
|||
QY 151 GATATTGGAAATGTATAAACAGGACAGGCTCTCTCGGACAGACAGAGAGGTTGAC 210
|||
Db 113 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 132
|||
QY 211 TTATCAAGAGTCTCCCTCACTCGAACAGCTTAACAGCAGATGAGAGTACTTCATCTCT 270
|||
Db 133 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgGlyPheIleSer 152
|||
QY 271 CACATCTTAGCCCTTTTTCGAGCGAGTGATCGAATTGTAATGAAATTTTGGTGGAGCGC 330
|||
Db 153 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 172
|||
QY 331 TTTAGTCAGAGGTGCGAGTTCCAGAGGCTCGCTGTTTCTATGGCTTTCAAATCTCATC 390
|||
Db 173 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 192
|||
QY 391 GAGAATGTTTCACTCAGAGATGATCAGTTTCTGATACAGACTTACATCAGAGATCCCAAG 450
|||
Db 193 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 212
|||
QY 451 AAAAGGGAATTTTATTATTAATGCAATTAATGCAATGCAATGCAATGCAATGCAATGCAAT 510
|||
Db 213 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 232
|||
QY 511 TGGGCTTGGCATGCGATGAGATAGAAATCTACTTTTGGGGAAGAGTGGTGGCCCTTT 570
|||
Db 233 TipAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaAlaPhe 252
|||
QY 571 GCTGCTGTAGAAGAGTTTCTTCTCAGGATCTTTTCTGCTGCTATATTCTGCTAAAGAG 630
|||
Db 253 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 272
|||
QY 631 AGAGGTCTTATGCCAGGACTCCTTTTTCAACTCAACTCAGCAGAGATGAAGACTTT 690
|||

```

Db 273 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuLeuSerArgAspGluGlyLeu 292
 QY CACTGTGACTTTGCTTGCCTGATGTTCCCAATCTTAGTAATAAGCCTTCAGAGAAGG 750
 Db 293 HisCysAspPheAlaCysLeuMetPheLeuHisLeuValHisLysProSerGluGluArg 312
 QY 751 GTCAGGGAGATCATTTCTGATGCTGTCACAAATTTGAGCAGGAGTTTAAACAGAACGCTTG 810
 Db 313 ValArgGluIleIleIleAsnAlaValArgIleGluGluPheLeuThrGluAlaLeu 332
 QY 811 CCAGTTGGCCTCATTCGAATGATTCATTTGATGAACAGTACATTTGAGTTTGTAGCT 870
 Db 333 ProValLeuLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 352
 QY 871 GACAGATTACTTGTGAATCTGATCTCAAAAGTTTTCAGGCACAAATCTCTTTGAT 930
 Db 353 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 372
 QY 931 TTTATGGAACAACTTTCTTTAGAGGAAACAAATTTCTTTGAGAAACGAGTTTCAGAG 990
 Db 373 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 392
 QY 991 TATCAGCTTTGTCAGTTATGTCAGAAACACAGATAAGCTCTTCACTTGGATGCAGAT 1050
 Db 393 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 412
 QY 1051 TTT 1053
 Db 413 Phe 413
 RESULT 15
 AAU28017
 ID AAU28017 standard; protein; 453 AA.
 AC AAU28017;
 XX
 DT 18-DEC-2001 (first entry)
 DE Human contig polypeptide sequence #170.
 XX
 KW Mammal; human; rhesus monkey; baker's yeast; fission yeast; Norway rat;
 KW mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;
 KW cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;
 KW nervous system disorder; inflammatory disorder; cell differentiation;
 KW angiogenesis; stem cell growth factor; activin; inhibin; cartilage; burn;
 KW genetic disorder; bone regeneration; tendon; ligament; tissue repair;
 KW cytostatic; antirheumatic; antiarthritic; vulnary; antiinflammatory;
 KW antibacterial; immunosuppressive; vasotropic; antiparkinsonian;
 KW neuroprotective; osteopathic; antidiabetic; antiasthmatic; antiallergic;
 KW immunostimulant; analgesic; gene therapy.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN WO200164834-A2.
 XX
 PD 07-SEP-2001.
 XX
 PF 26-FEB-2001; 2001WO-US004926.
 XX
 PR 28-FEB-2000; 2000US-00515126.
 PR 18-MAY-2000; 2000US-00577409.
 PR 17-JUN-2000; 2000US-00597707.
 PR 14-JUL-2000; 2000US-00616807.
 PR 19-SEP-2000; 2000US-00664641.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;
 PI Drmanac R;
 XX
 XX WPI; 2001-589862/66.

DR N-PSDB; AAS44917.
 XX Novel polypeptides and nucleic acids obtained from cDNA libraries
 PT prepared from various human tissues, for diagnosis, treatment of cancer,
 PT neurological, inflammatory disorders and for use in arrays for detection.
 XX
 XX Claim 10; Page 146-147; 153pp; English.
 XX
 CC Sequences AAU27676-AAU28019 represent full-length polypeptides and contig
 CC polypeptides of the invention. The proteins and their associated DNA
 CC sequences are useful for the treatment, diagnosis and prevention of
 CC various types of disorder in a mammalian subject such as a human, dog,
 CC monkey, mouse, hamster or rat. The disorders include cancers such as
 CC leukaemia, lymphoma and neuroblastoma, autoimmune disorders such as
 CC multiple sclerosis, connective tissue disease, rheumatoid arthritis,
 CC diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system
 CC disorders such as Parkinson's disease, Alzheimer's disease, Huntington's
 CC chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and
 CC Wernicke disease, inflammatory disorders such as nephritis, Crohn's
 CC disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory
 CC bowel disease. The sequences exhibit activity relating to angiogenesis,
 CC cell proliferation, cell differentiation, stem cell growth factor,
 CC activin or inhibin. Therefore, they can be used to manipulate stem cells
 CC in culture to give rise to neuroepithelial cells that can be used to
 CC augment or replace cells damaged by illness, accidental damage or genetic
 CC disorders. The sequences may also be used for regeneration of bone,
 CC cartilage, tendons and ligaments and in tissue repair and burn healing.
 CC Note: Some sequences for this patent did not form part of the printed
 CC specification, but were obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 453 AA;
 Alignment Scores:
 Pred. No.: 1.53e-165 Length: 453
 Score: 1454.00 Matches: 269
 Percent Similarity: 94.08% Conservative: 33
 Best Local Similarity: 83.80% Mismatches: 19
 Query Match: 78.76% Indels: 0
 DB: Gaps: 0
 US-10-698-228-12 (1-1053) x AAU28017 (1-453)
 QY 91 GAAGAGCCACTCTTAAGAAAGAGTCTCTCGCGGTTTGTCTTCATCTTCCATTCAGTACCT 150
 Db 133 AspGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGluTyrHis 152
 QY 151 GATATTGGAAAAATGTATAACAGGCACAGCTTCTCTTGGACAGCAGAGAGGTTGAC 210
 Db 153 AspIleTrpMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 172
 QY 211 TTATCAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAGTACTTCTCTCT 270
 Db 173 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 192
 QY 271 CACATCTTAGCCTTTTTCAGCCAGTGTGGAATTTAAATGAAATTTGGTGAGCGC 330
 Db 193 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 212
 QY 331 TTTAGTCAGGAGGTTCAGGTTCCAGAGCTCGCTGTTTCTATGGCTTTCAAATTTCTCATC 390
 Db 213 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 232
 QY 391 GAGAATGTTTCACTCAGAGATGTACAGTTTGTCTGATGACACTTTACATCAGAGATCCCAG 450
 Db 233 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 252
 QY 451 AAAAGGGAATTTTATTTAATGCAATTTGAAACCATGCCCTTATGTTAAGAAAAAGCAGAT 510
 Db 253 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysLysAlaAsp 272
 QY 511 TGGGCTTCGGATGGATAGCAGATAGAAATCTACTTTTGGGGAACAGTGTGCGCTTT 570

Db 273 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValValAlaPhe 292
QY 571 GCTGCTGTAGAGGAGTTTCTTCTCAGGATCTTTTGGCTATATCTTGGCTTAAAGAG 630
Db 293 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 312
QY 631 AGAGGTCTTATGCCAGGACTCACTTTTCCAAATGAATCATCAGCAGAGATGAAGGACTT 690
Db 313 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 332
QY 691 CACTGTGACTTTGCTTGCCTGATGTTCCTCAATACTTAGTAAATAAGCCTTCAGAAGAAAG 750
Db 333 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 352
QY 751 GTCAGGAGATCATTTGTTGATGCTGTCAAAATTGACGAGGAGTTTTCACAGAAGCCTTG 810
Db 353 ValArgGluIleIleIleAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 372
QY 811 CCAGTTGGCCTCATTTGGAATGAATTCATTTTCATGAAACAGTACATTCAGTTTGTAGCT 870
Db 373 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 392
QY 871 GACAGATTACTTGTGGAACTTGGATTCTCAAGGTTTTCAGGCAGAAAATCCTTTTGAT 930
Db 393 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 412
QY 931 TTTATGGAAAACATTTCTTAGAAGGAAAAACAATTTCTTTCAGAAACGAGTTTCAGAG 990
Db 413 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 432
QY 991 TATCAGCGTTTTCAGCTTATGGCAGAAACACAGATAACGTCCTTCACCTTGGATGCAGAT 1050
Db 433 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 452
QY 1051 TTT 1053
Db 453 Phe 453

Search completed: October 30, 2005, 06:28:52
Job time : 73.25 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: October 30, 2005, 05:15:40 ; Search time 15.75 Seconds
(without alignments)
12865.556 Million cell updates/sec

Title: US-10-698-228-12
Perfect score: 1846
Sequence: 1 atggcgaccggaaggcc.....tcacctggatgcagatttt 1053

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10698228/runat_26102005_100608_3995/app_query.fasta_1.2446
-DB=PIR_79 -QFMT=fastan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pcpt -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10698228@cgn_1_1_39 @runat_26102005_100608_3995 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR_79:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1821	98.6	366	2 T46249	hypothetical prote
2	1454	78.8	389	2 S25854	ribonucleoside-dip
3	1443	78.2	390	1 S06735	ribonucleoside-dip
4	1421	77.0	386	2 S27153	ribonucleoside-dip
5	1331	72.1	384	2 S24585	ribonucleoside-dip
6	1280	69.3	319	2 T30782	ribonucleoside-dip
7	1280	69.3	319	2 B72154	E4L protein - vari
8	1279	69.3	319	1 RUVZVV	ribonucleoside-dip
9	1279	69.3	333	2 T28466	ribonucleoside-dip
10	1266	68.6	333	2 H36839	C8L protein - vari
11	1258.5	68.2	348	2 T30470	ribonucleoside-dip
12	1217	65.9	299	1 ROSS2R	ribonucleoside-dip
13	1214	65.8	381	2 T18876	hypothetical prote
14	1169.5	63.4	399	1 A26916	ribonucleoside-dip

15	1161.5	62.9	391	2 T39992	ribonucleoside-dip
16	1161.5	62.9	391	2 S34808	ribonucleoside-dip
17	1128.5	61.1	329	2 T03688	ribonucleoside-dip
18	1070.5	58.0	340	2 S68538	ribonucleoside-dip
19	1024.5	55.5	349	2 B49412	ribonucleoside-dip
20	950	51.5	324	2 T17978	probable ribonucle
21	805	43.6	345	2 S59744	ribonucleoside-dip
22	681	36.9	668	2 T29884	hypothetical prote
23	638	34.6	327	1 RDVZAS	ribonucleoside-dip
24	385	20.9	331	2 A84389	ribonucleoside red
25	384.5	20.8	415	2 A83502	ribonucleoside red
26	358.5	19.4	346	2 G81728	ribonucleoside-dip
27	357.5	19.4	346	2 E71466	probable ribonucle
28	335	18.1	346	2 B72010	ribonucleoside-dip
29	335	18.1	346	2 F86613	ribonucleoside red
30	329	17.8	306	1 WMBE18	ribonucleoside-dip
31	324	17.6	303	2 T47526	ribonucleoside-dip
32	322	17.4	320	2 T42563	ribonucleoside-dip
33	318	17.2	321	1 WMBE11	ribonucleoside-dip
34	302.5	16.4	314	1 WMBE14	ribonucleoside-dip
35	301	16.3	302	1 WMBE12	ribonucleoside-dip
36	294.5	16.0	347	2 C87281	hypothetical prote
37	292.5	15.8	340	1 WMBE17	ribonucleoside-dip
38	290	15.7	305	2 S55655	ribonucleoside-dip
39	290	15.7	328	2 C97781	hypothetical prote
40	289.5	15.7	337	1 WMBE32	ribonucleoside-dip
41	288.5	15.6	337	1 WMBE12	ribonucleoside-dip
42	288.5	15.6	341	2 C71858	ribonucleoside-dip
43	286.5	15.5	341	2 D64565	ribonucleoside-dip
44	285	15.4	305	2 T03155	ribonucleoside-dip
45	285	15.4	324	2 B71655	ribonucleoside-dip

ALIGNMENTS

RESULT 1

T46249
hypothetical protein DKFZp761E1312.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C;Accession: T46249
R;Ansoorge, W.; Winkner, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A;Reference number: Z23028
A;Accession: T46249
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-366 <AAA>
A;Cross-references: UNIPROT:Q9NTD8; EMBL:AL137348
A;Experimental source: adult amygdala; clone DKFZp761E1312
C;Genetics:
A;Note: DKFZp761E1312.1
C;Superfamily: ribonucleoside reductase small subunit

Alignment Scores:
Pred. No.: 1.57e-157 Length: 366
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.65% Indels: 0
DB: 2 Gaps: 0

US-10-698-228-12 (1-1053) x T46249 (1-366)

QY	1	ATGGGCGACCCGGAAGCGCGCGCTGGATCAGATGAGATCATCTTCA	60
DB	16	MeUGlyaspProGluargProGluAlaAlaGlyLeuaspGlnaspGluargSerSer	35
QY	61	GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTCTAAGAAAGATTCTGC	120
DB	36	AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg	55
QY	121	CGGTTTGTTCATCTTTCCATCCAGTACCTCCGATATTGGAANAATGTATAAACAGCACAG	180

```
Db 56 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 75
QY 181 GCTTCTCTCTGACAGACAGAGGTTGACTTATCAAGAGTCTCCTCCTCCTGGAACAG 240
Db 76 AlaSerPheTrpThrAlaGluGluValAlaSerLysAspLeuProHisTrpAsnLys 95
QY 241 CTTAAAGCAGATGAGAACTTCTATCTCTCACATCTTAGCCCTTTTTCGACGACGTGAT 300
Db 96 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 115
QY 301 GGAATTGTAATGAAATTTGGTGGAGCCCTTTAGTCAGAGGTGCGAGTTCCAGAGGCT 360
Db 116 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 135
QY 361 CGCTGTTCTATGGCTTTCAAATCTCATCGAGAGTTCCTCATCAGAGATGACAGTTTG 420
Db 136 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 155
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAATTTTATTTAATGCAATTGAA 480
Db 156 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 175
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATGGCCCTTCGGATGGATAGCAGATAGAAA 540
Db 176 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 195
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTCGTGTAGAGGAGTTCCTTCTCAGGA 600
Db 196 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 215
QY 601 TCTTTTGTCTATATCTTGGCTTAAAGAGAGAGGTCTTATGCCAGAGCTCACTTTTCC 660
Db 216 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 235
QY 661 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTTGTCTTGGCTGATGTTCCAA 720
Db 236 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 255
QY 721 TACTTAGTAATAAGCCTTCAGAAAGAGGTGAGGAGATCATCTTGTGATGCTGTCAAA 780
Db 256 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 275
QY 781 ATTGACGAGCTTTTACAGAGCCTTGCAGTGGCTCATCTGGAATGCAATTCATT 840
Db 276 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 295
QY 841 TTGATGAAACAGTACATTGAGTTGTAGCTGCAGACTTACTGTGAACTTGTGAACTTCTCA 900
Db 296 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 315
QY 901 AAGTTTTTTCAGCAGAAAAATCCTTTTGTATTTATGGAATAACATTTCCTTAGAGAAAA 960
Db 316 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluLysLys 335
QY 961 ACAATTTCTTCAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC 1020
Db 336 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 355
QY 1021 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1053
Db 356 ThrAspAsnValPheThrLeuAspAlaAspPhe 366
```

RESULT 2

S25854
ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain - human
N:Alternate names: ribonucleotide reductase M2 chain; ribonucleotide reductase small chain
C:Species: Homo sapiens (man)
C:Date: 08-Jun-1994 #Sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C:Accession: S25854
R:Pavloff, N.; Rivard, D.; Masson, S.; Shen, S.H.; Mes-Masson, A.M.
DNA Seq. 2, 227-234, 1992
A:Title: Sequence analysis of the large and small subunits of human ribonucleotide reductase

A:Reference number: S25853; MUID:92329977; PMID:1627826

A:Accession: S25854

A:Molecule type: mRNA

A:Residues: 1-389 <P>

A:Cross-references: UNIPROT:P31350; EMBL:X59618; NID:G36154; PIDN:CAA42181.1; PID:G36155

C:Genetics:

A:Gene: GDB:RRM2

A:Cross-references: GDB:120358; OMIM:180390

A:Map position: 2p25-2p24

C:Superfamily: ribonucleoside reductase small subunit

C:Keywords: deoxyribonucleotide biosynthesis; iron; metalloprotein; oxidoreductase

F:138,169,172,233,266,269/Binding site: 2Fe-O cluster (Asp, Glu, His, Glu, His) #st

F:176/Active site: Tyr (stable tyrosyl radical) #status predicted

Alignment Scores:

Pred. No.:	4,1e-124	Length:	389
Score:	1454.00	Matches:	269
Percent Similarity:	94.08%	Conservative:	33
Best Local Similarity:	83.80%	Mismatches:	19
Query Match:	78.76%	Indels:	0
DB:	2	Gaps:	0

US-10-698-228-12 (1-1053) x S25854 (1-389)

```
QY 91 GAAGAGCCACTCTCTAAGAAAGAGTTCTCGCGGTTTGTTCATCTTCCAAATCCAGTACCCT 150
Db 69 AspGluProLeuLeuArgGluAsnProArgArgPheValIlePheProIleGluTyrHis 88
QY 151 GATATTGGAAATGATATAAACAGGACAGGCTTCTCTGGACAGCAGAGAGAGGTTGAC 210
Db 89 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 108
QY 211 TTATCAAGAGATCTCCCTCACTCGAACACAGCTTAAACAGACATGAGAGTACTTCATCT 270
Db 109 LeuSerLysAspIleGlnHisTrpGluSerLysProGluGluArgTyrPheIleSer 128
QY 271 CACATCTTAGCCCTTTTTCAGCCAGTGGATTTGTAATGAAATTTTCGTGGAGCGC 330
Db 129 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148
QY 331 TTTAGTCAGAGGTGACAGGTCCTCAGAGCTCGCTGTTTCTATGCTTTCATAATTTCTCAT 390
Db 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
QY 391 GAGAAATGTTCACTCAGAGATGTACAGTTGCTGTATAGACACTTACATCAGAGATCCCAAG 450
Db 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
QY 451 AAAAGGGAATTTTATTATGCAATGCAATGCAACCATGCCCTATGTTAAGAAAAAGCAGAT 510
Db 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysLysAlaAsp 208
QY 511 TGGGCTTCGAGTGGATAGCAGATAGAAAATCTACTTTTGGGAAAGAGTGGTGGCCTTT 570
Db 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaAlaPhe 228
QY 571 GCTGCTGTAGAAGAGTTTCTCTCAGAGATCTTTTCTGCTATATTTCTGCTATAATTTGCT 630
Db 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 248
QY 631 AGAGTCTTATGCCAGGACTCACCTTTTCCAACTCAATCATCAGCAGAGATGAAGACTT 690
Db 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
QY 691 CACTGTGACTTTTCTGCTGCTGATGTTTCCAATCTTAGTAATTAAGCCCTTCAGAGAAAGG 750
Db 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
QY 751 GTCAGGAGATCATTTGTTGATGTGTCGTCAAAATTCAGCAGAGGATTTTAAACAGAGCCTTG 810
Db 289 ValArgGluIleIleIleAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 308
QY 811 CCAGTTGGCCTCATTTGGAATGAATGTCATTTTGATGAACAGTACATTTGAGTTGTAGCT 870
```

```

Db      309 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 328
QY      871 GACAGATTACTTGTGGAACCTTGATTTCTCAAGGTTTTCAGGCAGAAATCCCTTTTGAT 930
Db      329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348
QY      931 TTTATGGAAAAACATTTCTTTAGAGGAAAAACAAATTTCTTTTCAGAAAAACAGTTTCAGAG 990
Db      349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368
QY      991 TATCAGCGTTTTCAGTTATGGCAGAAACACAGATAACGCTTTCACCTTGGATGCAGAT 1050
Db      369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
QY      1051 TTT 1053
Db      389 Phe 389

RESULT 3
S06735
ribonucleoside-diphosphate reductase (EC 1.17.4.1) chain M2 - mouse
N:Alternate names: ribonucleotide reductase M2 subunit
C:Species: Mus musculus (house mouse)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S06735; A24835
R:Thelander, M.; Thelander, L.
EMBO J. 8, 2475-2479, 1989
A:Title: Molecular cloning and expression of the functional gene encoding the M2 subunit
A:Reference number: S06735; MUID:90060004; PMID:2684652
A:Accession: S06735
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-390 <UNP>
A:Cross-references: UNIPROT:P11157; EMBL:X15666; NID:g50719; PIDN:CAA33707.1; PID:g50720
R:Thelander, L.; Berg, P.
Mol. Cell. Biol. 6, 3433-3442, 1986
A:Title: Isolation and characterization of expressible cDNA clones encoding the M1 and M2 subunits
A:Reference number: A24835; MUID:87089677; PMID:3025593
A:Accession: A24835
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-390 <GB>
A:Cross-references: GB:M14223; NID:g200767; PIDN:AAA40062.1; PID:g200768
A:Note: the authors translated the codon GGG for residue 315 as Glu
C:Genetics:
A:Map position: 12
A:Introns: 33/3; 58/3; 107/3; 146/3; 191/2; 223/1; 267/3; 302/3; 340/3
C:Superfamily: ribonucleoside reductase small subunit
C:Keywords: iron; oxidoreductase
F:139,170,173,233,267,270/Binding site: 2Fe-O cluster (Asp, Glu, His, Glu, His) #st
F:177/Active site: Tyr (stable tyrosyl radical) #status predicted

Alignment Scores:
Pred. No.: 4, 12e-123 Length: 390
Score: 1443.00 Matches: 269
Percent Similarity: 90.62% Conservative: 40
Best Local Similarity: 78.89% Mismatches: 26
Query Match: 78.17% Indels: 6
DB: 1 Gaps: 1

US-10-698-228-12 (1-1053) x S06735 (1-390)

QY      49 AGATCATCTTCAGACACCAAGAGTGAATAAAGTCA-----AAT 90
Db      50 ArgIlePheGlnAspSerAlaGluLeuGluSerLysAlaProThrAsnProSerValGlu 69
QY      91 GNAGAGCCACTCTTGAAGAGAGTTCTCGCCGGTTTGTCATCTTCCAAATCCAGTACCTT 150
Db      70 AspGluProLeuArgGluAsnProArgPheValPheProIleGluTyrHis 89
QY      151 GATATTTCGAAATCTATAACAGGCACACAGGCTTCTCTTCGACAGCAGAGGTTGAC 210

```

```

Db      90 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheThrTrpThrAlaGluGluValAsp 109
QY      211 TTATCAAGGATCTCCCTCACTGGAACAAGCTTAAGCAGATGAGAAGTACTTCTCTCT 270
Db      110 LeuSerLysAspIleGlnHisTrpGluAlaLeuLysProAspGluArgHisPheIleSer 129
QY      271 CATCATCTTAGCCTTTTTCAGCCAGTCATGGAATTTGAAATGAAAAATTTTGGTGGAGCCG 330
Db      130 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 149
QY      331 TTTAGTCAGAGGTGTCAGGTTTCAGAGCTCGCTGTTTCTATGCTTCAAAATTCATC 390
Db      150 PheSerGlnGluValGlnValThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 169
QY      391 GAGAACTTTCACACAGAGATGTACAGTTTGTGATAGACACTTACATCAGAGATCCCAG 450
Db      170 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 189
QY      451 AAAAGGGAATTTTATTTAATGCAATTGAAACCATGCCCTATGTTAAAGAAAAAGCAGAT 510
Db      190 GluArgGluTyrLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 209
QY      511 TGGCCCTTGCATGATAGCAGATAGAAAATCTACTTTTGGGAAAGAGTGGTGCCTTT 570
Db      210 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaPhe 229
QY      571 GCTGCTGTAGAAGAGTTTCTTCTCAGGATCTTTTGTGCTGCTATATTCTGGCTAAAGAAG 630
Db      230 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 249
QY      631 AGAGGTCTTATGTCAGGACTCACATTTTCCAAATCACTCATCAGCAGAGATGAAGACTT 690
Db      250 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluLeu 269
QY      691 CACTGTGACTTTGCTGCTGATGTTCCAACTACTTAGTAAATAAGCCTTCAGAGAAAAGG 750
Db      270 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProAlaGluGlnArg 289
QY      751 GTCAGGAGATCATTTGTTGATGCTGTCATAAATTGAGCAGAGAGTTTAAACAGAGCCTTG 810
Db      290 ValArgGluIleIleThrAsnAlaValArgIleGluGlnPheLeuThrGluAlaLeu 309
QY      811 CCAGTTGGCCTCATTTGGAATGAAATGCAATTTGATGAAACAGTACATTGAGTTTGTAGCT 870
Db      310 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 329
QY      871 GACAGATTACTTGTGGAACCTTGATTTCTCAAGGTTTTCAGGCAGAAAAATCCCTTTGAT 930
Db      330 AspArgLeuMetLeuGluLeuGlyPheAsnLysIlePheArgValGluAsnProPheAsp 349
QY      931 TTTATGGAAAAACATTTCTTTAGAGGAAAAACAAATTTCTTTTCAGAAACGAGTTTCAGAG 990
Db      350 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 369
QY      991 TATCAGCGTTTTCAGTTATGGCAGAAACACAGATAACGCTTTCACCTTGGATGCAGAT 1050
Db      370 TyrGlnArgMetGlyValMetSerAsnSerThrGluAsnSerPheThrLeuAspAlaAsp 389
QY      1051 TTT 1053
Db      390 Phe 390

RESULT 4
S27153
ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain - golden hamster
N:Alternate names: ribonucleotide reductase small chain
C:Species: Mesocricetus auratus (golden hamster)
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C:Accession: S27153
R:Chaudhuri, M.M.; Tonin, P.N.; Srinivasan, P.R.
Biochim. Biophys. Acta 1171, 117-121, 1992
A:Title: cDNA sequence of the small subunit of the hamster ribonucleotide reductase.
A:Reference number: S27153; MUID:93042000; PMID:1384717

```

4

Db 199 AlaAspTrpAlaMetArgTrpIleAsnAspAspSerSerSerTyrAlaGluArgValVal 218
 QY 565 GCCTTGTCTGTAGAACAGATTTCTCTCAGGATCTTTGCTGCTATATCTCGCTA 624
 Db 219 AlaPheAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeu 238
 QY 625 AGAAGAGAGGCTTATGCGCAGGACTCCTCTTTTCCAACTCATCAGCAGAGATGAA 684
 Db 239 LysLysArgGlyIleMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGlu 258
 QY 685 GGACTTCATCTGACTTTCCTGCTGCTGATGTTCCCAATACCTAGTAAATAGCCCTCAGAA 744
 Db 259 GlyLeuHisCysAspPheAlaCysLeuMetPheSerHisLeuValAsnLysProSerGln 278
 QY 745 GAAAGGCTCAGGAGATCATTTGATGCTGTGCAAAATTTGACGAGAGTTTAAACAGAA 804
 Db 279 GluArgIleHisGlnIleIleAspGluAlaValLysIleGlnValPheLeuThrGlu 298
 QY 805 GCCTTGCCAGTTGGCCTCATTTGGAATGGAATTCATTTGATGAAACAGTACATTTGAGTTT 864
 Db 299 AlaLeuProCysArgLeuIleGlyMetAsnCysAspLeuMetArgGlnTyrIleGluPhe 318
 QY 865 GTAGCTGACAGATTTACTTGTGGAATCTCAAGTTCTCAAGTTTTCAGGCAGAAATCTCT 924
 Db 319 ValAlaAspArgLeuLeuLeuLysCysAspLysLeuTyrAsnLysGluAsnPro 338
 QY 925 TTTGATTTTATGAAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACAGATT 984
 Db 339 PheAspPheMetGluHisIleSerLeuGluGlyLysThrAsnPhePheGluLysArgVal 358
 QY 985 TCAGAGTATCAGGTTTTCAGTTATGCA-----GAAACACACAGAT---AACGCTCTC 1035
 Db 359 GlyGluTyrGlnLysMetGlyValMetSerGlyGlyAsnThrGlyAspSerHisAlaPhe 378
 QY 1036 ACCTTGATGTCAGATTTT 1053
 Db 379 ThrLeuAspAlaAspPhe 384
 RESULT 6
 T30782
 Ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain - vaccinia virus (strain
 N1) Alternate names: ribonucleotide reductase, small subunit
 C:Species: vaccinia virus
 A:Variety: strain Ankara
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 R:Antoine, G.; Scheiflinger, F.; Falkner, F.G.; Dorner, P.
 Submitted to the EMBL Data Library, March 1997
 A:Description: The complete genomic sequence of the Modified Vaccinia Ankara (MVA) strain
 A:Reference number: Z20877
 A:Accession: T30782
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-319 <ANT>
 A:Cross-references: UNIPROT:O57175; EMBL:U94848; PIDN:AB96415.1
 A:Experimental source: strain Ankara
 C:Genetics:
 A:Note: MVA032L
 C:Function:
 A:Description: catalyzes the reduction of ribonucleoside diphosphate to deoxyribonucleoside
 C:Superfamily: ribonucleoside reductase small subunit
 C:Keywords: deoxyribonucleotide biosynthesis; oxidoreductase
 Alignment Scores:
 Pred. NO.: 2.73e-108 Length: 319
 Score: 1280.00 Matches: 239
 Percent Similarity: 85.31% Conservative: 34
 Best Local Similarity: 74.69% Mismatches: 45
 Query Match: 69.34% Indels: 2
 DB: 2 Gaps: 2

US-10-698-228-12 (1-1053) x T30782 (1-319)

QY 94 GAGCCACTCCTCAAGAAAGAGTTCTCGCCGGTTTGTCTCATCTTTTCAATCCAGTACCCCTGAT 153
 Db 2 GluProIleLeuAlaProAsnArgPheValIlePheProIleGlnTyrHisAsp 21
 QY 154 ATTTGGAAATATGATATAACAGGACAGGCTTCTCTTCTGGACAGCAGAGAGAGGTTGACTTA 213
 Db 22 IleTrpAsnMetTyrLysLysAlaGluAlaSerPheTrpThrValGluGluValAspIle 41
 QY 214 TCAAGAGATCTCCCTCACTCGAAACAAGCTTAAACAGCAGATGAGAGTACTTCTCATCTCTCAC 273
 Db 42 SerLysAspIleAsnAspTrpAsnLysLeuThrProAspGluLysTyrPheLysHis 61
 QY 274 ATCTTAGCCCTTTTTCAGCCAGTGTGAATTTGTAATGAAATTTTGGTGGAGCGCTTT 333
 Db 62 ValLeuAlaPhePheAlaAlaSerAspGlyLysValAsnGluAsnLeuAlaGluArgPhe 81
 QY 334 AGTCAGAGAGTGCAGGTTCCAGAGGCTCGCTGTTTCTATGGCTTTTCAAAATTTCTCATCAG 393
 Db 82 CysThrGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnMetAlaIleGlu 101
 QY 394 AATGTTCACTCAGAGATGTACAGTTTCTGCTGATAGACACTTACATCAGAGATCCCAAGAAA 453
 Db 102 AsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrValLysAspSerAsnGlu 121
 QY 454 AGGGAATTTTATTAATGCAATTGAAACCATGCCCTATGTTAAGAAAAAGAGCAGATTGG 513
 Db 122 LysAsnTyrLeuPheAsnAlaIleGluThrMetProCysValLysLysLysAlaAspTrp 141
 QY 514 GCCTTGCAGTGTAGATGAGATGAGAAATCTACTTTTGGGAAAGAGTGTGGCTTCTGCT 573
 Db 142 AlaGlnLysTrpIleHisAsp---SerAlaGlyTyrGlyGluArgLeuIleAlaPheAla 160
 QY 574 GCTGTAGAAGAGATTTTCTCTCAGGATCTTTGCTGCTATATTTCTGGCTAAAGAGAGA 633
 Db 161 AlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLysArg 180
 QY 634 GGTCTTATGCCAGGACTCAGTTTTCCTCAATGAACTCATCAGCAGAGATGAAGGACTTCAC 693
 Db 181 GlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHis 200
 QY 694 TGTGACTTCTCTGCTGATGTTTCCAACTACTTAAATAAGCCTTCAGAGAAAGGGTCT 753
 Db 201 CysAspPheAlaCysLeuMetPheLysHisLeuHisProProSerGluGluThrVal 220
 QY 754 AGGAGATCATTTGTTGATGCTGCTCAAAATTTGAGCAGAGATTTTAAACAGAGCTTGCCA 813
 Db 221 ArgSerIleIleThrAspAlaValSerIleGluGlnGluPheLeuThrAlaAlaLeuPro 240
 QY 814 GTTGGCTCATTTGGAATGATTCATTTTTCATGAAACAGTACATTTGAGTTTGTAGCTGAC 873
 Db 241 ValLysLeuIleGlyMetAsnCysGluMetLysThrTyrIleGluPheValAlaAsp 260
 QY 874 AGATTACTTGTGGAACCTTGATTTCTCAAGGTTTTCAGGCAGAGAAATTCCTTTTGTATTT 933
 Db 261 ArgLeuIleSerGluLeuGlyPheLysLysIleTyrAsnValThrAsnProPheAspPhe 280
 QY 934 ATGGAACACATTTCTTTTAGAAGGAAACAAATTTCTTTGAGAAACGAGTTTTCAGAGTAT 993
 Db 281 MetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGluTyr 300
 QY 994 CAGCGTTTTCAGTTATGCGAGAAACCCACAGATACCTTTCACCTTGGATGCGAGATTTT 1053
 Db 301 GlnLysMetGlyValMetSerGln---GluAspAsnHisPheSerLeuAspValAspPhe 319

RESULT 7

B72154
E4L protein - variola minor virus (strain Garcia-1966)

C:Species: variola minor virus

C:Date: 24-Nov-1999 #sequence_revision 24-Nov-1999 #text_change 09-Jul-2004

C:Accession: B72154

R:Shchelkunov, S.N.; Totmenin, A.V.; Gutorov, V.V.; Safronov, P.F.; Massung, R.F.; Lopar

submitted to GenBank, March 1998

A:Description: Analysis of the complete coding sequence of DNA of alastrim variola minor

A;Reference number: A72150
A;Accession: B72154
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-319 <SHC>
A;Cross-references: UNIPROT:Q89087; GB:Y16780; NID:G5830555; PIDN:CA54628.1; PID:G5830555
A;Experimental source: strain Garcia-1966
C;Genetics:
A;Gene: E4L
C;Superfamily: ribonucleoside reductase small subunit

Alignment Scores:
Pred. No.: 2,73e-108 Length: 319
Score: 1280.00 Matches: 240
Percent Similarity: 85.31% Conservative: 33
Best Local Similarity: 75.00% Mismatches: 45
Query Match: 69.34% Indels: 2
DB: 2 Gaps: 2

US-10-698-228-12 (1-1053) x B72154 (1-319)

QY 94 GAGCCACTCTAAGAAAGAGTTCGCCGGTTGTCTATCTTCCATCCAGTACCTCGAT 153
Db 2 GluProIleuAlaLysAsnProAsnArgPheValIlePheProIleGlnTyrHisAsp 21
QY 154 ATTTGGAAATGTATAACAGGCACAGCGTCTCTCTGGACAGCAGAGGTTGACTTA 213
Db 22 IleTrpAsnMetTyrLysLysAlaGluAlaSerPheTrpThrValGluGluValAspIle 41
QY 214 TCAAGGATCTCCCTCACTGGAACAGCTTAAGCAGATGAGAGTACTCTCTCAC 273
Db 42 SerLysAspIleAsnAspTrpAsnLysLeuThrProAspGluLysTyrPheIleLysHis 61
QY 274 ATCTTAGCTTTTTCACCGCAGTGTGAATGTAAATGTAATTTGGTGAGCGCTTT 333
Db 62 ValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuAlaGluArgPhe 81
QY 334 AGTCAGAGCTGAGGTTCCAGAGGCTCGCTGTTCTATGGCTTTCAAATCTCATCGAG 393
Db 82 CysIleGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnMetAlaIleGlu 101
QY 394 AATGTTCACTCAGAGATGACAGTGTTCGTGATAGACACTTACATCAGAGATCCCAAGAA 453
Db 102 AsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrValLysAspSerAsnGlu 121
QY 454 AGGGAATTTTATTAAATGCAATGGAACATGCCCTATGTTAAGAAAAAAGCAGATTGG 513
Db 122 LysAsnTyrLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAspTrp 141
QY 514 GCCTTGCAGATGATACAGATAAATCTACTTTTGGGGAAGAGTGGTGGCTTTGCT 573
Db 142 AlaGlnLysTrpIleHisAsp---SerAlaGlyTyrGlyGluArgLeuIleAlaPheAla 160
QY 574 GCTGTAGAAGAGTATTTCTCTCAGATCTTTGCTGTATATCTGGCTAAAGAGAGA 633
Db 161 AlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLysArg 180
QY 634 GGTCTTATGCCAGGACTACTTTTCCAAATGAACACTCATCAGCAGAGATGAAGACTTCAC 693
Db 181 GlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHis 200
QY 694 TGTGACTTTCCTGCTGATGTTCCTAATCTAGTAAATAAGCTTCAGAGAAGGGTC 753
Db 201 CysAspPheAlaCysLeuMetPheLysHisLeuLeuTyrProProSerGluGluThrVal 220
QY 754 AGGAGATCATTTGTCATGCTGTCAAAATTGACAGAGTGTTTTAAACAGAGCTTGCCA 813
Db 221 ArgSerIleIleThrAspAlaValSerIleGluGlnGluPheLeuThrValAlaLeuPro 240
QY 814 GTTGCGCTCATTTGGAATGATTTGCAATTTTGAAGAAACAGTACATTTGATGTGAC 873
Db 241 ValLysLeuIleGlyMetAsnCysGluMetMetLysThrTyrIleGluPheValAlaAsp 260

QY 874 AGATTACTTGTGGAACTTGATTTCTCAAGAGTGTTCAGCAGAGAAAAATCCTTTTGATTTT 933
Db 261 ArgLeuIleSerGluLeuGlyPheLysLysIleTyrAsnValThrAsnProPheAspPhe 280
QY 934 ATCGAAAAATTTCTTTAGAGAAAAAACHAATTTCTTTTGGAGAAACGAGTTTCAGATAT 993
Db 281 MetGluAsnIleSerLeuGluGlyThrAsnPhePheGluLysArgValGlyGluTyr 300
QY 994 CAGCGTTTTCAGTTATGCGAGAAACACACATAAAGCTCTTACCTTGGATGCAGATTTT 1053
Db 301 GlnLysMetGlyValMetSerGln---GluAspAsnHisPheSerLeuAspValAspPhe 319

RESULT 8
RDVZVV
ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain - vaccinia virus
N;Alternate names: F4L protein
C;Species: vaccinia virus
C;Date: 30-Jun-1989 #sequence revision 30-Jun-1989 #text_change 09-Jul-2004
C;Accession: A29892; I42506; F36213
R;Slabaugh, M.; Roseman, N.; Davis, R.; Mathews, C.
J. Virol. 62, 519-527, 1988
A;Title: Vaccinia virus-encoded ribonucleotide reductase: sequence conservation of the
A;Reference number: A29892; MUID:88091062; PMID:2826813
A;Accession: A29892
A;Molecule type: DNA
A;Residues: 1-319 <SLA>
A;Cross-references: UNIPROT:P11158; GB:M19117; NID:G335808; PIDN:AAA88680.1; PID:G335809
A;Experimental source: strain WR
R;Goebel, S.J.; Johnson, G.P.; Perkus, M.E.; Davis, S.W.; Winslow, J.P.; Paoletti, E.
Virology 179, 517-563, 1990
A;Title: Appendix to "The complete DNA sequence of vaccinia virus".
A;Reference number: A42501
A;Accession: I42506
A;Molecule type: DNA
A;Residues: 1-212, 'Y', 214-319 <GOB>
A;Cross-references: GB:M35027; NID:G335317; PIDN:AAA48018.1; PID:G335366
A;Experimental source: strain Copenhagen
R;Goebel, S.J.; Johnson, G.P.; Perkus, M.E.; Davis, S.W.; Winslow, J.P.; Paoletti, E.
Virology 179, 247-266, 1990
A;Title: The complete DNA sequence of vaccinia virus.
A;Reference number: A42531; MUID:91021027; PMID:2219722
A;Contents: annotation; possible protein-coding frames
A;Note: neither amino acid nor nucleotide sequence is given
R;Roseman, N.A.; Slabaugh, M.B.
Virology 178, 410-418, 1990
A;Title: The vaccinia virus HindIII F fragment: nucleotide sequence of the left 6.2 kb.
A;Reference number: A36213; MUID:91020979; PMID:2219701
A;Accession: F36213
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-319 <ROS>
A;Cross-references: EMBL:M34368; NID:G335618; PIDN:AAA48244.1; PID:G335624
A;Experimental source: strain WR
C;Function:
A;Description: catalyzes the reduction of ribonucleoside diphosphate to deoxyribonucleoside
C;Superfamily: ribonucleoside reductase small subunit
C;Keywords: DNA replication; iron; metalloprotein; oxidoreductase
F;70,104,163,197,200/Binding site: 2Fe-O cluster (Asp, Glu, His, Glu, His) #sta
F;108/Active site: Tyr (stable tyrosyl radical) #status predicted

Alignment Scores:
Pred. No.: 3,36e-108 Length: 319
Score: 1279.00 Matches: 239
Percent Similarity: 85.31% Conservative: 34
Best Local Similarity: 74.69% Mismatches: 45
Query Match: 69.28% Indels: 2
DB: 1 Gaps: 2

US-10-698-228-12 (1-1053) x RDVZVV (1-319)

QY 94 GAGCCACTCTAAGAAAGAGTTCGCCGGTTGTCTATCTTCCATCCAGTACCTCGAT 153
Db 2 GluProIleuAlaLysAsnProAsnArgPheValIlePheProIleGlnTyrHisAsp 21

Db 275 ArgLeuIleSerGluLeuGlyPheLysLysIleTyrAsnValThrAsnProPheAspPhe 294
QY 934 ATGGAAACATTTCTTAGAAGGAGAAACAAATTTCTTTGAGAAACGAGTTTCAGAGTAT 993
Db 295 MetGluAsnIleSerLeuGluGlyLysThrAsnProPheGluLysArgValGlyGluTyr 314
QY 994 CAGCGTTTGGCAGTTATGCGACAGAACCCAGATAACGCTCTTCACCTTGGATGAGATTTT 1053
Db 315 GlnLysMetGlyValMetSerGln---GluAspAsnHisPheSerLeuAspValAspPhe 333

RESULT 10
H36839
CBL protein - variola virus (strain India-1967)
C:Species: variola virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C:Accession: H36839
R:Blinov, V.M.
submitted to GenBank, November 1992
A:Reference number: A36859
A:Accession: H36839
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-333 <BLI>
A:Cross-references: UNIPROT:P33799; GB:X69198; NID:G456758; PIDN:CAA48969.1; PID:G297209
C:Superfamily: ribonucleoside reductase small subunit

Alignment Scores:
Pred. No.: 5.18e-107 Length: 333
Score: 1266.00 Matches: 237
Percent Similarity: 85.00% Conservative: 35
Best Local Similarity: 74.06% Mismatches: 46
Query Match: 68.58% Indels: 2
DB: 2 Gaps: 2

US-10-698-228-12 (1-1053) x H36839 (1-333)

QY 94 GAGCCACTCTTAAGAAAGAGTTCTCGCCGGTTGTCTCATCTTTCCAAATCCAGTACCTCGAT 153
Db 16 GluProIleLeuAlaLysAsnProAsnArgPheValIlePheProIleGlnTyrHisAsp 35
QY 154 ATTTGGAAATGTATAACAGGCACAGGCTTCTCTTGGACACGACGAGAGGTTTGACTTA 213
Db 36 IleTrpAsnMetTyrLysLysAlaGluAlaSerPheTrpThrValGluGluValAspIle 55
QY 214 TCAAGGATCTCCCTCACTGGAACAGCTTAAGACAGATGAGAGTACTTCACTCTCAC 273
Db 56 SerLysAspIleAsnAspTrpAsnLysLeuThrProAspGluLysTyrPheIleLysHis 75
QY 274 ATCTTAGCCTTTTTCAGCCAGTGTGAATGTAAATGAAATTTGGTGGAGCGCTTT 333
Db 76 ValLeuAlaPhePheAlaAlaSerArgPheValIleValAsnGluAsnLeuAlaGluArgPhe 95
QY 334 AGTCAGGAGGTGCAGGTTCCAGAGGTCGCTGTTCTTATGGCTTTCAAAATTTCTCATCGAG 393
Db 96 CysIleGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnMetAlaIleGlu 115
QY 394 AATGTTCACTCAGAGATGTACAGTTGCTGATAGACACTTACATCAGATCCCAAGAA 453
Db 116 AsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrValLysSerSerAsnGlu 135
QY 454 AGGGAATTTTATTAATGCAATTTGAACCATCCCTATGTTTAAGAAAGAAAGCAGAGTTGG 513
Db 136 LysAsnTyrLeuPheAsnAlaIleGluThrMetProCysValLysLysLysAlaAspTrp 155
QY 514 GCCTTGCAGTAGATAGACAGATAAATCTACTTTTGGGAAAGAGAGTGGCTTTGGCT 573
Db 156 AlaGlnLysTrpIleHisAsp---SerAlaGlyTyrGlyGluArgLeuIleAlaPheAla 174
QY 574 GCTGTAGAGGAGTTTCTTCTCAGAGATCTTTTGGCTGCTATATCTTGGCTAAAGAGAGA 633
Db 175 AlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLysArg 194
QY 634 GGTCTTATCCAGGACTCACTTTTCCATGAATCACTCATCAGCAGAGATGAAGACTTCAC 693

Db 195 GlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHis 214
QY 694 TGTGACTTTCCTGCTGATGTTCCCAATACTTACTAGTAATAAGCCCTTCAGAGAAGGCTC 753
Db 215 CysAspPheAlaCysLeuMetPheLysHisLeuLeuTyrProProSerGluGluThrVal 234
QY 754 AGGAGATCATCTGTGATGCTGCAAAATTCAGCAGGAGTTTAAACAGAGCCCTTGCCA 813
Db 235 ArgSerIleIleThrAspAlaValSerIleGluGlnGluPheLeuThrValAlaLeuPro 254
QY 814 GTTGGCCTCATTCGAATGAATTCATTTGATGAAACAGTACATTGAGTTTGTAGCTGAC 873
Db 255 ValLysLeuIleGlyMetAsnCysGluMetMetLysThrTyrMetGluPheValAlaAsp 274
QY 874 AGATTACTTCTGGAACCTTGATCTCAAGGTTTTCAGGCAGAGAAATCCCTTTGATTTT 933
Db 275 ArgLeuIleSerGluLeuGlyPheLysArgIleTyrAsnValThrAsnProSerAspPhe 294
QY 934 ATGGAAACATTTCTTTAGAAAGGAAACAAATTTCTTTGAGAAACGAGTTTTCAGAGTAT 993
Db 295 MetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGluTyr 314
QY 994 CAGCGTTTTCAGTTATGCGACAGAACCCAGATAACGCTCTTCACTTGGATGCGAGATTT 1053
Db 315 GlnLysMetGlyValMetSerGln---GluAspAsnHisPheSerLeuAspValAspPhe 333

RESULT 11
T30470
ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain - Lymantria dispar nuclea
N:Alternate names: ribonucleotide reductase small subunit homolog
C:Species: Lymantria dispar nuclear polyhedrosis virus, LdMNPV
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T30470
R:Kuzio, J.; Pearson, M.N.; Harwood, S.H.; Funk, C.J.; Evans, J.T.; Slavicek, J.M.; Rohr
Virology 253, 17-34, 1999
A:Title: Sequence and analysis of the genome of a baculovirus pathogenic for Lymantria d
A:Reference number: Z20836; MUID:99124785; PMID:9887315
A:Accession: T30470
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-348 <KUZ>
A:Cross-references: UNIPROT:Q9YMK7; EMBL:AF081810; PIDN:AAC70306.1
C:Function:
A:Description: catalyzes the reduction of ribonucleoside diphosphate to deoxyribonucleos
C:Superfamily: ribonucleoside reductase small subunit
C:Keywords: oxidoreductase; pyrimidine deoxynucleotide metabolism

Alignment Scores:
Pred. No.: 2.52e-106 Length: 348
Score: 1258.50 Matches: 246
Percent Similarity: 80.63% Conservative: 37
Best Local Similarity: 70.09% Mismatches: 59
Query Match: 68.17% Indels: 9
DB: 2 Gaps: 3

US-10-698-228-12 (1-1053) x T30470 (1-348)

QY 1 ATGGCGCACCCGGAAGAGCCGCGGCGGCTGATCAGGATGAGAGATCATCTTCA 60
Db 7 MetProAlaProGluArgPro-----PheAspProArgAlaProAla 20
QY 61 GACACCAACGAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 21 Pro-----GluArgProPheAspProArgAlaGluProLeuLeuArgGluAsnProArg 38
QY 121 CGGTTTCTCATCTTCCCAATCCAGTACCTCATATTTGGAAAAATGTATAACAGGCACAC 180
Db 39 ArgPheValIlePheProIleGlnTyrProAspMetTrpArgMetTyrLysLysAlaGlu 58
QY 181 GCTTCTCTTGGACAGCAGAGAGGTTGACTTATCAAGAGGATCTCCCTCCTCAGTGGACAC 240
Db 59 AlaSerPheTrpThrValGluGluValAspLeuSerLysAspThrSerAspTrpGluArg 78

C;Function:

A;Description: catalyzes the reduction of ribonucleoside diphosphate to deoxyribonucleosides

C;Superfamily: ribonucleoside reductase small subunit

C;Keywords: deoxyribonucleotide biosynthesis; iron; metalloprotein; oxidoreductase

F;45,76,139,173,176/Binding site: 2Fe-O cluster (Asp, Glu, His, Glu, Glu, His) #status

F;83/Active site: Tyr (stable tyrosyl radical) #status predicted

Alignment Scores:

Pred. NO.:	1.46e-102	Length:	299
Score:	1217.00	Matches:	233
Percent Similarity:	88.67%	Conservative:	33
Best Local Similarity:	77.67%	Mismatches:	30
Query Match:	65.93%	Indels:	4
DB:	1	Gaps:	3

US-10-698-228-12 (1-1053) x RDS2R (1-299)

Qy	163	ATGATATAACAGGCACAGCGTTCCTCTCGACAGCAGAGAGGTGTGACTATCAAGGAT	222
Db	1	MetTyrLysAlaGluAlaSerPheTrp--AlaGluGluValAspLeuSerLysAsp	19
Qy	223	CTCCCTCACTGAAACAAGCTTAAGACAGATGAGAAGTACTTCACTCTCACATCTTAGCC	282
Db	20	MetAlaHisTrpGluSerLeuLysGluGluLysHisPheSerHisValLeuAla	39
Qy	283	TTTTTTCAGCCAGCATGATGGAAATTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAG	342
Db	40	PhePheAlaAlaSerAspGlyLeValAsnGluAsnLeuValGluArgPheSerLysGlu	59
Qy	343	GTGAGGTTTCAGAGGTCGCTGTTTCTATGGCTTTCAAAATTCATCGAGAAATGTTTCAAC	402
Db	60	ValGlyValThrGluAlaArgCysPheTyrGlyPheGlnLeAlaMetGluAsnIleHis	79
Qy	403	TCAGAGATGACAGTTTGCTGATAGACACTTACATCAGAGATCCCAGAGAAAGGAATTT	462
Db	80	SerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProGlnGluArgAspPhe	99
Qy	463	TTATTTTATGCAATTGAAACCACTGCGCTATGTTTAAAGAAAAGCAGATTTGGCCTTGCGA	522
Db	100	LeuPheAsnAlaIleGluThrMetProCysValLysGluLysAlaAspTrpAlaMetArg	119
Qy	523	TGGATAGCAGATAGAAAATCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAA	582
Db	120	TrpIleAsnAspSerSerTyrAlaGluArgValAlaPheAlaAlaValGlu	139
Qy	583	GGAGTTTCTCTCAGGATCTTTCTGCTATATCTGCTAAGAGAGAGAGGCTTATG	642
Db	140	GlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLysArgGlyIleMet	159
Qy	643	CCAGGACTCACTTTTCCAAATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTT	702
Db	160	ProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPhe	179
Qy	703	GCTTGCCTGATGTTTCCAATACCTAGTAAATAAGCCTTCAGAGAAAGGGTCAGGAGATC	762
Db	180	AlaCysLeuMetPheSerHisLeuValAsnLysProSerGlnGluArgIleHisGlnIle	199
Qy	763	ATTGTTGATGCTGTCAAAATTTGAGCAGGAGTGTATTAACAGAGCCTTGCCAGTTGGCCTC	822
Db	200	IleAspGluAlaValLysIleGluGlnValPheLeuThrGluAlaLeuProCysArgLeu	219
Qy	823	ATTGGAATGAATTGCATTTTGGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTT	882
Db	220	IleGlyMetAsnCysAspLeuMetArgGlnTyrIleGluPheValAlaAspArgLeuLeu	239
Qy	883	GTGGAACCTTGGATTCTCAAGGTTTTTTCAGGCAGAGAAATCTTTTTGATTTTTATCGAAAC	942
Db	240	LeuGluLeuLysCysAspLysLeuTyrAsnLysGluAsnProPheAspPheMetGluHis	259
Qy	943	ATTTCTTTTAGAGGAAAAACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTT	1002
Db	260	IleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGluTyrGlnLysMet	279

Db 218 IleAlaPheAlaIaValGluGlyIlePheSerGlySerPheAlaSerIlePheTrp 237

QY 622 CTAAGAGAGAGAGGTCTTATGCAGAGACTCACTTTTCCAAATGAACCTCATCAGCAGAGAT 681
|||||
Db 238 LeuLysLysArgGlyLeuMetProGlyLeuThrHisSerAsnGluLeuIleSerArgAsp 257
|||||
QY 682 GAAGGACTTCACGTGACCTTGCTCGCTGGCTGATTGCCAATACCTTAGTAAATAAGCCTTCA 741
|||||
Db 258 GluGlyLeuHisArgAspPheAlaCysLeuLeuTyrSerLysLeuGlnLysLysLeuThr 277
|||||
QY 742 GAAGAAGGGTCAGGAGACATCTGTGATGCTGTCAAATTCAGCAGGAGTGTAAAAACA 801
|||||
Db 278 GlnGlnArgIleTyrAspIleLeuLysAspAlaValIleGluGlnGluPheLeuThr 297
|||||
QY 802 GAAGCCTTGCAGTTCGGCTCATTTGGAATGAATGCTTTTGTGATGAACAGTACATTCAG 861
|||||
Db 298 GluAlaLeuProValaspMetIleGlyMetAsnCysArgLeuMetSerGlnTyrIleGlu 317
|||||
QY 862 TTTGTAGCTGACAGATTACTTGTGAACTTGGAATCTCAAAGTTTTTTCAGGCAGAAAAT 921
|||||
Db 318 PheValAlaAspHisLeuLeuValGluLeuGlyCysAspLysLeuTyrLysSerLysAsn 337
|||||
QY 922 CCTTTTCAGATTTATGGAACAACTTTCTTTAGAGGAGAAAAACAAATTTCTTTTGAGAACGA 981
|||||
Db 338 ProPheaspPheMetGluAsnIleSerIleAspGlyLysThrAsnPhePheGluLysArg 357
|||||
QY 982 GTTTCAGAGTATCACGCGTTTTCAGTTATGGCAGAAACCACACATACGCTCTTCACCTTG 1041
|||||
Db 358 ValSerGluTyrGlnArgProGlyValMetValAsnGluAlaGluArgGlnPheAspLeu 377
|||||
QY 1042 GATGCAGATTTT 1053
:::
Db 378 GluAlaAspPhe 381
|||||

RESULT 14

A26916

ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein J1271; protein YJL026w; ribonucleotide reductase small chain
C;Species: Saccharomyces cerevisiae
C;Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: A26916; S07605; S56798
R;Elledge, S.J.; Davis, R.W.
Mol. Cell. Biol. 7, 2783-2793, 1987
A;Title: Identification and isolation of the gene encoding the small subunit of ribonucleotide reductase
A;Reference number: A26916; MUID:88038817; PMID:3313004
A;Accession: A26916
A;Molecule type: DNA
A;Residues: 1-399 <ELL>
A;Cross-references: UNIPROT:P09938; EMBL:M17221
A;Note: The authors translated the codon GAA for residue 101 as Lys, GCT for residue 102 as Ser, GCT for residue 107 as Phe, and GAA for residue 108 as Trp
A;Note: the sequence shown follows the authors' translation
R;Hurd, H.K.; Roberts, C.W.; Roberts, J.W.
Mol. Cell. Biol. 7, 3673-3677, 1987
A;Title: Identification of the gene for the yeast ribonucleotide reductase small subunit
A;Reference number: S07605; MUID:88065506; PMID:3316984
A;Accession: S07605
A;Molecule type: DNA
A;Residues: 1-399 <HUP>
A;Cross-references: EMBL:M17789; NID:g172449; PIDN:AAA34988.1; PID:g172450
R;Pohl, T.M.; Aljinovic, G.
submitted to the Protein Sequence Database, September 1995
A;Reference number: S56793
A;Accession: S56798
A;Molecule type: DNA
A;Residues: 1-399 <TOV>
A;Cross-references: EMBL:Z49301; NID:g1008141; PIDN:CAA89317.1; PID:g1008142; GSPDB:GN000142
C;Genetics:
A;Gene: SGD,RNR2; MIPS:YJL026w
A;Cross-references: SGD:S0003563; MIPS:YJL026w
A;Map position: 10L
C;Function:
A;Description: oxidoreductase; pyrimidine deoxynucleotide metabolism

C:Superfamily: ribonucleoside reductase small subunit

C:Keywords: iron; oxidoreductase; pyrimidine deoxynucleotide metabolism
F:145,176,179,239,273,276/Binding site: 2Fe-O cluster (Asp, Glu, His, Glu, His) #sd
F:183/Active site: Tyr (stable tyrosyl radical) #status predicted

Alignment Scores:

Pred. No.: 3 286-98 Length: 399
Score: 1169.50 Matches: 221
Percent Similarity: 81.07% Conservative: 53
Best Local Similarity: 65.36% Mismatches: 58
Query Match: 63.35% Indels: 6
DB: 1 Gaps: 3

US-10-698-228-12 (1-1053) x A26916 (1-399)

```

QY 54 ATCTTCAGACCAACCAAGTGAATAAAGTCAATGAAGCCACTCTTAAGAAGAG 113
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 64 VAHISARGHISLysLeuLys-GluMetGlu---LysGluGluProLeuLeuAsnGluAs 82
QY 114 TTCTGCGCGGTTGTTCATCTTTCCAAATCCAGTACCTGTATTTGGAAATGTATAACA 173
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 82 physGluArgThrValLeuPheProIleLysTyrHisGluIleTrpGlnAlaTyrLysAr 102
QY 174 GGCACAGCGCTTCTTCTGACAGCAGAGAGGTGTACTTATCAAGAGATCTCCCTCAGTG 233
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 102 gAlaGluAlaSerPheTrpThrAlaGluIleAspLeuSerLysAspIleHisAspTr 122
QY 234 G---NACAAGCTTAAGCAGATGAGAGTACTTCTCATCTTCATCTTACCTTTTTCG 290
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 122 pAsnAsnArgMetAsnGluAsnGluArgPheIleSerArgValLeuAlaPhePheAl 142
QY 291 AGCCAGTATGAAATTTGTAATAATTTGTGAGCGCTTTAGTCAGAGGTGCGAGT 350
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 142 alalaserAspGlyIleValAsnGluAsnLeuValGluAsnPheSerThrGluValGlnI 162
QY 351 TCACAGAGCTCGCTGTTCTATGGCTTTTCAAAATTTCTCATCGAGAATGTTCACCTCAGAGAT 410
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 162 eProGluAlaLysSerPheTyrGlyPheGlnIleMetIleGluAsnIleHisSerGluTh 182
QY 411 GTACAGTTGCTGTATGACACTTACATCAGACATCCCAAGAAAGGAAATTTTATTAA 470
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 182 rTyrSerLeuLeuIleAspThrTyrIleLysAspProLysGluSerGluPheLeuPheAs 202
QY 471 TGCATTGAACACCTGCTTATTTAAGAAAGCAGATTCGGCTTCGGATGAGTAC 530
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 202 nAlaIleHisThrIleProGluIleGlyGluLysAlaGluTrpAlaLeuArgTrpIleGl 222
QY 531 AGATGAAATCTACTTTTGGGAAAGAGTGTGGCTTTGTGCTGTAGAGGAGCTTTT 590
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 222 nAspAlaAspAlaLeuPheGlyGluArgLeuValAlaPheAlaSerIleGluGlyValPh 242
QY 591 CTCTCAGAGATCTTTTGTCTATATTCGGTAAAGAGAGAGGTCTTTATGCCAGGACT 650
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 242 ePheSerGlySerPheAlaSerIlePheTrpLeuLysLysArgGlyMetMetProGlyLe 262
QY 651 CACTTTTTCATGATCACTCATCAGCAGATGAAGACTTCACTGTGACTTTTGTTCGCT 710
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 262 uThrPheSerAsnGluLeuIleCysArgAspGluGlyLeuHisThrAspPheAlaCysLe 282
QY 711 GATGTTTCAATCTTAGTAAATAAGCTTTTCAAGAAAGGTCAGGAGATCATTTGTGA 770
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 282 uLeuPheAlaHisLeuLysAsnLysProAspProAlaIleValGluLysIleValThrGl 302
QY 771 TGCTGTCAAAATTTGACGAGGAGTTTATTAACAGAGCCTTGGCCTTCATTGGGAAT 830
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 302 uAlaValGluIleGluGlnArgTyrPheLeuAspAlaLeuProValAlaLeuLeuGlyWe 322
QY 831 GAATTGCATTTTGTATGAACAGTACATTGAGTTTGTAGTGTGACAGATTACTTGTGAAC 890
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 322 tAsnAlaAspLeuMetAsnGlnTyrValGluPheValAlaAspArgLeuLeuValAlaPh 342
QY 891 TGGATTCTCAAGGTTTTCAGGCGAATAATCTTTGATTTTATGGAAACATTTCTTT 950
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

Db 342 eGlyAsnLysLysTyrTyrLysValGluAsnProPheAspPheMetGluAsnIleSerLe 362
QY 951 AGAAGGAAACAAATTTCTTTCAGAAACGAGTTTCAGAGTATCAGCGCTTTTGCAGTTAT 1010
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 362 uAlaGlyLysThrAsnPhePheGluLysArgValSerAspTyrGlnLysAlaGlyValMe 382
QY 1011 GGCAGAAACCCACAGAT-----AACGTCTTCCACTTGGATGTCAGATTTT 1053
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 382 tSerLysSerThrLysGlnGluAlaGlyAlaPheThrPheAsnGluAspPhe 399

RESULT 15

T39992

ribonucleoside-diphosphate reductase small chain - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C:Accession: T39992

R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Whithead, S.; Chillingworth, T.; Churcher,

submitted to the EMBL Data Library, July 1998

A:Reference number: Z21897

A:Accession: T39992

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-391 <LYN>

A:Cross-references: UNIPROT:P36603; EMBL:AL031158; PIDN:CAA20100.1; GSPDB:GN00067; SPDB:

A:Experimental source: strain 972h-; cosmid c25D12

C:Genetics:

A:Gene: SPDB:SPBC25D12.04

A:Map position: 2

C:Superfamily: ribonucleoside reductase small subunit

Alignment Scores:

Pred. No.: 1 75e-97 Length: 391
Score: 1161.50 Matches: 227
Percent Similarity: 78.80% Conservative: 48
Best Local Similarity: 65.04% Mismatches: 61
Query Match: 62.92% Indels: 13
DB: 2 Gaps: 4

US-10-698-228-12 (1-1053) x T39992 (1-391)

```

QY 37 GATCAGATGAGAGATCATCTTCAGAC-----ACCAACGAAAGTGAATAAAG 84
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 46 GluGluLysGlnGluGlyAspTyrTyrLeuGlyLysLysGluAspGluLeu--- 64
QY 85 TCAATGAAGAGCCACTCTTAAGAAAGAGTCTCGCGTTTGTTCATCTTCCATCCAG 144
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 65 -----AspGluValValLeuArgProAsnProHisArgPheValLeuPheProIleLys 82
QY 145 TACCCTGATATTGGAAATGTATAACAGCAGCAGGCTCTCTCGACAGCAGAGAG 204
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 83 TyrHisGluIleTrpGlnPheTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGlu 102
QY 205 GTTGACTTATCAAGGATCTCCCTCACTGG---AACCAAGCTTAAAGCAGATGAGAGTAC 261
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 103 IleAspLeuSerLysAspLeuValAspTrpAspAsnLysLeuAsnAlaAspGluArgTyr 122
QY 262 TTCACTCTCACAATCTAGCTTTTTCAGCCAGTGTGAATGTAAATGAAATTTG 321
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 123 PheIleSerThrValLeuAlaTyrPheAlaSerAspGlyIleValAsnGluAsnLeu 142
QY 322 GTGGAGCGCTTACTCAGGAGGTGCAGGTTCAGAGGCTCGCTGTTTCTATGGCTTCAA 381
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 143 LeuGluArgPheSerSerGluValGlnIleProGluAlaArgCysValTyrGlyPheGln 162
QY 382 ATTCTCATTCAGAAATGTTCCTACTCAGAGATGTACAGTTTGTGTAGACACATTCATCAGA 441
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 163 IleMetIleGluAsnIleHisSerGluThrTyrSerLeuLeuLeuAspThrTyrIleArg 182
QY 442 GATCCCAAGAAAGGAAATTTTATTAAATGCAATGAAACCATGCCCTATGTTAAGAAA 501
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 183 GluProLysGluLysGlnArgHisPheAspAlaIleLeuThrMetGlySerIleLysAla 202
QY 502 AAAGCAGATTGGCGCTTGGATGCGATGATAGATAGAAAATCTACTTTTGGGGAAGAGTG 561

```


GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 30, 2005, 03:28:20 ; Search time 76 Seconds
(without alignments)

14189.984 Million cell updates/sec

Title: US-10-698-228-12

Perfect score: 1846

Sequence: 1 atggcgaccggaaagcc.....tcaccttgatcgagatttt 1053

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 3224756

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p.model -DEV=xlh
-Q/cgn2.1/USPTO.spool/US10698228/runat.26102005.100608.3981/app.query.fasta.1.2446
-DB=Uniprot_03 -QFMT=fastan -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=PCT -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10698228 @CGN 1.1.149 @runat.26102005.100608.3981 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Uniprot_03.*

1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1821	98.6	351	2 Q7LG56	Q7LG56 homo sapien
2	1817	98.4	351	2 Q9NUW3	Q9NUW3 homo sapien
3	1725	93.4	351	2 Q6PEE3	Q6PEE3 mus musculus
4	1534	83.1	297	2 Q6A141	Q6A141 homo sapien
5	1514	82.0	299	2 Q7SPY9	Q7SPY9 homo sapien
6	1474	78.8	285	2 Q86YE3	Q86YE3 homo sapien
7	1454	78.8	389	1 R1R2_HUMAN	P11350 homo sapien
8	1443	78.2	390	1 R1R2_MOUSE	P11157 mus musculus
9	1439	78.0	386	2 Q68E1	Q68E1 xenopus tro
10	1431	77.5	386	2 Q801Q4	Q801Q4 xenopus lae
11	1427	77.3	386	2 Q8AVY2	Q8AVY2 xenopus lae
12	1425.5	77.2	386	1 R1R2_BRARE	P79733 brachydanio
13	1421	77.0	386	1 R1R2_MESAU	Q60561 mesocricetu
14	1420.5	77.0	386	2 Q6D144	Q6D144 brachydanio
15	1399.5	75.8	378	2 Q6P876	Q6P876 xenopus tro
16	1385	75.0	406	2 Q61P47	Q61P47 xenopus lae

ALIGNMENTS

RESULT 1

Q7LG56 PRELIMINARY; PRT; 351 AA.
ID Q7LG56 Q9NPD6; Q9NTD8;
AC Q7LG56; Q9NPD6; Q9NTD8; 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 28, Last annotation update)
DE Ribonucleotide reductase (Hypothetical protein DKFZp761E1312) (P53-
inducible ribonucleotide reductase small subunit 2).
GN Name=p53R2; Synonyms=DKFZp761E1312;
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20179179; PubMed=10716435; DOI=10.1038/35003506;
RA Tanaka H., Arakawa H., Yamaguchi T., Shiraishi K., Fukuda S.,
RA Matsui K., Takei Y., Nakamura Y.;
RT "A ribonucleotide reductase gene involved in a p53-dependent cell-
cycle checkpoint for DNA damage.";
RL Nature 404:42-49(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Amalgam;
RG The German cDNA Consortium;
RA Ansoorge W., Krieger S., Regiert T., Rittmuller C., Schwager B.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Ugai H., Yokoyama K.K.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB036532; BAA92493.1; JOINED.
DR EMBL; AB036524; BAA92493.1; JOINED.
DR EMBL; AB036525; BAA92493.1; JOINED.
DR EMBL; AB036526; BAA92493.1; JOINED.
DR EMBL; AB036527; BAA92493.1; JOINED.
DR EMBL; AB036528; BAA92493.1; JOINED.
DR EMBL; AB036529; BAA92493.1; JOINED.

Q7pf28 anopheles g
Q7qif4 anopheles g
Q7zyw0 brachydanio
Q95vp8 aedes aegypti
Q27124 urechis cau
P48592 drosophila
Q9xyn8 aedes albop
P07201 spigula sol
O6cfu6 yarowia li
O5175 vaccinia vi
Q76q46 variola min
Q89087 variola vir
P11158 vaccinia vi
Q76zx1 vaccinia vi
Q6zqg9 rabbitpox v
P20493 vaccinia vi
Q8jlh6 ectromelia
Q89559 variola vir
Q8gnl4 cowpox viru
P29883 vaccinia vi
Q8v544 monkeypox v
P87632 cowpox viru
Q87632 camelpox vi
Q77525 camelpox vi
Q91fe9 vaccinia vi
P31799 variola vir
Q9c167 neurospora
Q9pxr4 variola vir
Q9ymk7 lymantria d

DR EMBL; AB036530; BAA92493.1; JOINED.
 DR EMBL; AB036531; BAA92493.1; JOINED.
 DR EMBL; AL137348; CAB70703.2; -.
 DR EMBL; AB166671; BAD12267.1; -.
 DR EMBL; AB036063; BAA92434.1; -.
 DR Genew; HGNC:17296; RRM2B.
 DR GO; GO:0004748; P:ribonucleoside-diphosphate reductase activity; IEA.
 DR GO; GO:0009186; P:deoxyribonucleoside diphosphate metabolism; IEA.
 DR InterPro; IPR009078; Perritin/RR like.
 DR InterPro; IPR000358; Ribonuclei_redctase.
 DR Pfam; PF00268; Ribonuc_red sm; 1.
 DR PROSITE; PS00368; RIBRED_SMALL; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 351 AA; 40736 MW; 6D008687EEF40994 CRC64;

Alignment Scores:

Pred. No.: 3.18e-151 Length: 351
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 98.65% Indels: 0
 DB: 2 Gaps: 0

US-10-698-228-12 (1-1053) x Q7LG56 (1-351)

QY 1 ATGGCGCGCCGGAAGCGCGGAGCGGCGGCTGGATCAGATCAGATCATCTTCA 60
 DB 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
 QY 61 GACACCAAGCAAGTCAATTAAGTCAATTAAGTCAAGTCAAGTCAAGTCAAGTCA 120
 DB 21 AspThrAsnGluSerGluLeuLeuLeuSerAsnGluGluProLeuLeuArgSerSerArg 40
 QY 121 CGTTTGTTCATCTTCCAAATCCAGTACCTGATATTTGGAAATGATTAACAGGCACAG 180
 DB 41 ArgPheValIlePheProIleGlnTyProAspIleTrpLysMetTrpLysGlnAlaGln 60
 QY 181 GCTTCTCTCTGACACGACGAGGTTGACTTATCAAGGATCTCCTCCTCAGTCAAG 240
 DB 61 AlaSerPheTrpThrAlaGluGluValAlaAspLeuSerLysAspLeuProHisTrpAsnLys 80
 QY 241 CTTAAAGCAGATGAGAAGTACTCTCTCATCTCTTACATCTTACGCTTTTTCGAGCCAGTGAT 300
 DB 81 LeuLysAlaAspGluLysTyPheLeuSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
 QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGTTCAGAGGCT 360
 DB 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnProGluAla 120
 QY 361 CGCTGTTTCTATGGCTTTTCAATTTCTCATCGAGAAATGTTCACTCAGATGTACAGTTTG 420
 DB 121 ArgCysPheTyGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTySerLeu 140
 QY 421 CTGATAGACACTTACATCAGATGCCAAGAAAGGAAATTTTATTAATGCAATTTGAA 480
 DB 141 LeuIleAspThrTyIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCGCTTTCGATGGATAGCAGTAAAAA 540
 DB 161 ThrMetProTyValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
 QY 541 TCTACTTTGGGGAAGAGTGTGGCTTGTCTGCTGTGTGAGGAGTCTTCTCTCAGGA 600
 DB 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
 QY 601 TCTTTTGTGCTATATCTTGGCTTAAGAGAGAGAGGCTTATGCGGAGTCTCACTTTTTC 660
 DB 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 QY 661 AATGAACCTCATCAGCAGATCAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
 DB 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240

QY 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGCTCAGAGAGATCATTTGTGCTGTCTCAA 780
 DB 241 TyrlleuValAsnLysProSerGluGluArgValArgGluIleValAspAlaVallys 260
 QY 781 ATTGACGAGGAGTATTTAAACAGAGCCTTCCTCAGATTTGGCTCATTTGGATTCATT 840
 DB 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysile 280
 QY 841 TTGATGAACAGTACATAGTGTGTAGCTGCACAGATTAAGTGTGGAAGTTCGATTCCTCA 900
 DB 281 LeuMetLysGlnTyrlleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 QY 901 AAGTGTTCCTCAGCAGAAAAATCCCTTTGATTTTATGAAACATTTCTTTAGAAGAAAA 960
 DB 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlylys 320
 QY 961 ACAAATTTCTTCAGAAACAGTTCACAGTATCAGCGTTTTCGAGTTTATGCGAGAAACC 1020
 DB 321 ThrAsnPhePheGluLysArgValSerGluTyGlnArgPheAlaValMetAlaGluThr 340
 QY 1021 ACAGATAACGTCTTCACCTTGGATGCAGATTTT 1053
 DB 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 2

Q9NUW3 PRELIMINARY; PRT; 351 AA.
 AC Q9NUW3;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein FLJ11103.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX PubMed=14702039; DOI=10.1038/ngi1285;
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
 RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
 RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K.,
 RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
 RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
 RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
 RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
 RA Tanai H., Kimata M., Watanabe M., Hiraoaka S., Chiba Y., Ishida S.,
 RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,
 RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
 RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
 RA Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
 RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
 RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shiget K., Senba T.,
 RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
 RA Togaishi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
 RT "Complete sequencing and characterization of 21,243 full-length human
 cDNAs.";
 RL Nat. Genet. 36:40-45 (2004).
 DR EMBL; AK01965; BAA92005.1; -.
 DR HSSP; P11157; IXSM.

DR GO: 0004748; F: ribonucleoside-diphosphate reductase activity; IEA.
 DR GO: 0009186; P: deoxyribonucleoside diphosphate metabolism; IEA.
 DR InterPro: IPR009078; Ferritin/RR like.
 DR Ribonuc: RIB000358; Ribonuc_redctse.
 DR Pfam: PF00268; Ribonuc_red_sm; 1.
 DR PROSITE: PS00368; RIBRED_SMALL; 1.
 SQ SEQUENCE 351 AA; 40704 MW; 6D009B2D59E9A323 CRC64;

Alignment Scores:
 Pred. No.: 7, 15e-151 Length: 351
 Score: 1917.00 Matches: 350
 Percent Similarity: 100.00% Conservatives: 1
 Best Local Similarity: 99.72% Mismatches: 0
 Query Match: 98.43% Indels: 0
 DB: 2 Gaps: 0

US-10-698-228-12 (1-1053) x Q9NUW3 (1-351)

QY 1 ATGGGCGACCCGGAAGCGCGGCTGGATCAGGATCAGATCATCTTCA 60
 Db 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
 QY 61 GACACACGAAGTGAATAAGTCAATGAAGAGCCACTCTAAGAAAGTCTCCG 120
 Db 21 AspThrAsnGluSerGluLeuLeuSerAsnGluGluProLeuLeuArgLysSerArg 40
 QY 121 CGTTTGTTCATCTTCCATCCAGTACCTGATATTTGAAAATGTATAAAGCCACAG 180
 Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
 QY 181 GCTTCTCTCTGACAGCAGAGAGGTGCTACTTATCAAGGATCTCCCTCAGTGAACAG 240
 Db 61 AlaSerPheThrPheAlaGluGluValAspLeuSerLysAspLeuProIleThrPheLys 80
 QY 241 CTTAAAGCAGATGAGAAGTACTCTCTCATCTTTAGCCCTTTTTCGACCCAGTGTAT 300
 Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisLeuLeuAlaPhePheAlaAlaSerAsp 100
 QY 301 GGAATGTAAATGAATTTGTGAGGCTTTAGTACGAGGTGAGTCCAGAGGCT 360
 Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 QY 361 CGCTGTTCTATGGCTTCAATCTCTCATCGAAGATGTTCACTCAGAGATGACAGTTTG 420
 Db 121 ArgCysPheTyrGlyPheGlnIleLeuLeuGluAsnValHisSerGluMetTyrSerLeu 140
 QY 421 CTGATAGACACTTACATCAGATCCCAAGAAAGGGAATTTTATTAATCAATTGAA 480
 Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 QY 481 ACCATGCCCTATGTTAAGAAAGACAGATTGGGCTTGCATGATGATGATAGAAAA 540
 Db 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
 QY 541 TCTACTTTTGGGAAGAGTGGTGGCTTGTCTGTAGAGAGGATTTCTTCTCAGA 600
 Db 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
 QY 601 TCTTTTGTCTATATTCTGGCTAAAGAGAGAGTCTTATCCAGGAGCTCACTTTTCC 660
 Db 201 SerPheAlaAlaIlePheThrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 QY 661 AATGAACATCATCAGCAGATGAAGACTTCACTGTGACTTTGCTGTGCTGTGTTCCAA 720
 Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 QY 721 TACTTAGTAATAACCTTCAGAACAGGTCAGGAGATCATTTGATGCTGTCAA 780
 Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleValAspAlaValLys 260
 QY 781 ATTGAGCAGGAGTCTTTTAAACAGACCTTGGCCTTGCAGTTGGCTCATTTGGAATGCA 840
 Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyValAsnCysIle 280

QY 841 TTGATGAAAACAGTACATTTGAGTTTGTAGCTGCAGAGATTACTTGTGGAACTTGGATTCTCA 900
 Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 QY 901 AAGTTTTTTCAGGACAGAAAATCTTTTGTATTTATGAAACATTTCTTTAGAAAGGAAAA 960
 Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 QY 961 ACAATTTCTTTGAGAAACAGTTTCAGACTATCAGCGTTTTCAGTTATGCGCAGAACACC 1020
 Db 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 QY 1021 ACAGATAACCTCTTTCACCTTGGATGAGATTTT 1053
 Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 3

Q6PEE3 PRELIMINARY; PRT; 351 AA.
 AC O6PEE3; 05-JUL-2004 (T-EMBLrel. 27, Created)
 DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
 DE Ribonucleotide reductase M2 B (Tp53 inducible).
 GN Name=Rm2b;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RX MEDLINS=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton D., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Sklasko U., Small U., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC058103; AAH58103.1; -
 DR GO; GO:0004748; Ribonucleoside-diphosphate reductase activity; IEA.
 DR GO; GO:0009263; P: deoxyribonucleotide biosynthesis; IEA.
 DR InterPro; IPR009078; Ferritin/RR like.
 DR Ribonuc: RIB000358; Ribonuc_redctse.
 DR Pfam; PF00268; Ribonuc_red_sm; 1.
 DR PROSITE; PS00368; RIBRED_SMALL; 1.
 SQ SEQUENCE 351 AA; 40803 MW; 4E1259233C9CC8A9 CRC64;

Alignment Scores:
 Pred. No.: 8, 75e-143 Length: 351
 Score: 1725.00 Matches: 328
 Percent Similarity: 97.44% Conservatives: 14
 Best Local Similarity: 93.45% Mismatches: 9

Query Match:	93.45%	Indels:	0
DB:	2	Gaps:	0
US-10-698-228-12 (1-1053) x Q6PEB3 (1-351)			
QY	1	ATGGGCGACCCGGAAGCGCGGCTGGATCAGATGAGATCATCTTCA	60
Db	1	MetGlyAspProGluArgProGluAlaAaArgProGluLysGlyGluInLeuCysSer	20
QY	61	GACACCAAGCAAGTAAATGAATGAAGACCACTCTAAGAAGAGTTCTCCG	120
Db	21	GluThrGluGluAsnValValArgSerAsnGluProLeuLeuArgLysSerSerArg	40
QY	121	CGTTTGTCTATCTTCCATCCAGTACCTGATATTGAAAATGATATAACAGGCACAG	180
Db	41	ArgPheValIlePheProIleGlnTyrProAspIleTrpArgMetTyrLysGlnAlaGln	60
QY	181	GTTCTCTTCGACAGCAGAGAGGTGACTTATCAAGGATCTCCCTCACTGGAACAAG	240
Db	61	AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys	80
QY	241	CTTAAGCAGATGAGAGTACTTCTCATCTCTACATCTTAGCCCTTTTTCAGCCAGTAT	300
Db	81	LeuLysSerAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAaSerAsp	100
QY	301	CGAATTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT	360
Db	101	GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValProGluAla	120
QY	361	CGCTGTTTCTATGGCTTCAAAATTCATCGAAGTGTTCACATCAGAGATGTACAGTTTG	420
Db	121	ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu	140
QY	421	CTGATAGACACTTACATCAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA	480
Db	141	LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu	160
QY	481	ACCATGCCCTATGTTAAGAAAAAGCAGATTGGCCCTCGATGGATAGCAGATAGAAA	540
Db	161	ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys	180
QY	541	TCTACTTTTGGGGAAGAGTGTGGCTTTGCTGTAGAGGAGTTTCTTCTCAGGA	600
Db	181	SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyIlePhePheSerGly	200
QY	601	TCCTTTGCTGCTATATCTGGCTTAAAGAGAGAGGTCTTATGCGCAGACTCTCTTTTCC	660
Db	201	SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer	220
QY	661	AATGAATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA	720
Db	221	AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln	240
QY	721	TACTTAGTAAATAAGCCTTTCAGAAAGAGGTCAGGAGATCATTTGATGCTGTCAAA	780
Db	241	TyrLeuValAsnLysProSerGluAspArgValArgGluIleIleAlaAspAlaValGln	260
QY	781	ATTGACAGAGTTTAAACAGAGCCTTGCAGTGGCTCATTCGGAATGAATTCATT	840
Db	261	IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysVal	280
QY	841	TTGATGAACAGTACATTGAGTTGTAGCTGACAGATTACTTGTGAACCTTGATTTCTCA	900
Db	281	LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuGlyLeuGlyPheSer	300
QY	901	AAGGTTTTTCAGCAGAAAATCTTTTGTATTTTATGAAAAACATTTCTTTAGAGGAAA	960
Db	301	LysIlePheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys	320
QY	961	ACAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC	1020
Db	321	ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr	340
RESULT 4			
Q6AI41			
QY	1021	ACAGATAAGCTCTTACCTTGGATGAGATTTT 1053	
Db	341	ThrAspAsnValPheThrLeuAspAlaAspPhe 351	
PRELIMINARY; PRT; 297 AA.			
QY	1021	ACAGATAAGCTCTTACCTTGGATGAGATTTT 1053	
Db	341	ThrAspAsnValPheThrLeuAspAlaAspPhe 351	
ALIGNMENT SCORES:			
Pred. No.:	5,28e-126	Length:	297
Score:	1534.00	Matches:	296
Percent Similarity:	99.66%	Conservative:	0
Best Local Similarity:	99.66%	Mismatches:	1
Query Match:	83.10%	Indels:	0
DB:	2	Gaps:	0
US-10-698-228-12 (1-1053) x Q6AI41 (1-297)			
QY	163	ATGTATAACAGCAGCAGCTTCTTCGACAGCAGAGGTTGACTTATCAAGGAT	222
Db	1	MetTyrLysGlnAlaGlnAlaSerPheTrpThrAlaGluValAspLeuSerLysAsp	20
QY	223	CTCCCTCACTGGAACAGCTTAAAGCAGATGAGAAGTACTTCTCTCATCTTAGCC	282
Db	21	LeuProHisTrpAsnLysLeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAla	40
QY	283	TTTTTTGACGCCAGTGAATTTGTAATGAAATTTGTTGGAGCGCTTTTAGTCAGGAG	342
Db	41	PhePheAlaAlaSerAspGlyIleValAsnGluValGluArgPheSerGlnGlu	60
QY	343	GTCCAGGTTCCAGAGCGCTGCTGTTTCTATGGCTTTCAAATTTCTCATCGAATGTTAC	402
Db	61	ValGlnValProGluAlaArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHis	80
QY	403	TCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTT	462
Db	81	SerGluMetTyrSerLeuLeuIleAspThrTyrIleArgAspProLysLysArgGluPhe	100
QY	463	TTATTTAAATGCAATTCGAAACCATCGCCCTATGTTTAAAGAAAAAGCAGATTGGGCTTGC	522
Db	101	LeuPheAsnAlaIleGluThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArg	120
QY	523	TGATAGCAGATAGAAATCTACTTTTGGGAAAGAGTGTGGCTTTTGTCTGTAGAA	582
Db	121	TrpIleAlaAspArgLysSerThrPheGlyGluArgValValAlaPheAlaAlaValGlu	140
QY	583	GGAGTTTCTTCTCAGATCTTTTGTGCTATATTCTGGCTAAAGAGAGGCTTCTATG	642

```

Db      141 GlyValPheSerGlySerPheAlaAlaPheTrpLeuLysLysArgGlyLeuMet 160
QY      643 CCAGGACTCATTCTTCCCAATGAACTCATCAGCAGAGATGAAGGACTTCACTGACTTT 702
Db      161 ProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPhe 180
QY      703 GCTTGCTGATGTTCCTCAATCTAGTAAATAAGCCTTCAGAGAAAGGGTCAGGAGATC 762
Db      181 AlaCysLeuMetPheGlnTyrLeuValAsnLysProGluGluArgValArgGluLeu 200
QY      763 ATTGTTGATGCTGCTCAAAATGAGCAGGAGTCTTTTAAACAGAGCCTTGCCAGTTGGCCTC 822
Db      201 IleValAspAlaValLysIleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeu 220
QY      823 ATTGGAATCAATTCATTTGATGAACAGTACATTCAGTTGCTGAGTGCAGATTTACTTT 882
Db      221 IleGlyMetAsnCysIleLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeu 240
QY      883 GTGGAATCTGGATCTCAAGAGTCTTTCAGGCAGAAATCTTTGATTTTATGGAATAC 942
Db      241 ValGluLeuGlyPheSerLysValPheGlnAlaGluAsnProPheAspPheMetGluAsn 260
QY      943 ATTCTTTTGAAGAGAAACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTT 1002
Db      261 IleSerLeuGluGlyLysThrAsnPhePheGluLysArgValSerGluTyrGlnArgPhe 280
QY      1003 GCAGTTATGCCAGAAACACAGATACGCTTTCACCTTGATGCAGATTTT 1053
Db      281 AlaValMetAlaGluThrThrAspAsnValPheThrLeuAspAlaAspPhe 297

```

RESULT 5

```

Q75PY9 PRELIMINARY; PRT; 299 AA.
AC Q75PY9;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DE P53-inducible ribonucleotide reductase small subunit 2 long form.
GN Name=p53R2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RA Ugai H., Yokoyama K.K.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB163437; BAD11774.1; -.
DR GO; GO:0004748; P:ribonucleoside-diphosphate reductase activity; IEA.
DR GO; GO:0003186; P:deoxyribonucleoside diphosphate metabolism; IEA.
DR InterPro; IPR009078; F:Ferritin/RR like.
DR InterPro; IPR000358; R:ribonucleotide.
DR Pfam; PF00268; R:Ribonuc red sm; 1.
DR PROSITE; PS00368; R:RIBRED SMALL; 1.
SQ SEQUENCE 299 AA; 34528 MW; 4705C44389EB689B CRC64;

```

Alignment Scores:

```

Pred. No.: 3.03e-124 Length: 299
Score: 1514.00 Matches: 299
Percent Similarity: 85.19% Conservative: 0
Best Local Similarity: 85.19% Mismatches: 0
Query Match: 82.02% Indels: 52
DB: 2 Gaps: 1

```

US-10-698-228-12 (1-1053) x Q75PY9 (1-299)

```

QY      1 ATGGGGGACCCGGAAGAGCGCGGCTGGATCAGATGAGATCATCTTCA 60
Db      1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAsp----- 15
QY      61 GACACCAACGAAAGTGAATAAGTCAATGAAGCCACTCTCTAAGAAAGAGTTCTCGC 120

```

```

Db      15 ----- 15
QY      121 CGGTTTGTCACTCTTCCCAATCCAGTACCCTGATATTGTGAAATGTATAAACAGGCACAG 180
Db      15 ----- 15
QY      181 GCTTCTTCTGGACAGCAGAGAGAGGTTGACTTATCAAAAGGATCTCCCTCAGTGGAAACAG 240
Db      16 -----GluValAspLeuSerLysAspLeuProHisTrpAsnLys 28
QY      241 CTTAAACAGATGAGAAGTACTTCTCATCTCTCATCTTTAGCCTTTTTCAGCCAGTGTAT 300
Db      29 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 48
QY      301 GGAATCTTAATGAAATTTGGTCGAGCGCTTTAGTCAGGAGGTCACAGGTTCCAGAGGCT 360
Db      49 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 68
QY      361 CGCTGTTTCTATGGCTTTCAAAATCTCATCCAGAAATGTTCTACTCAGAGATGTACAGTTTG 420
Db      69 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 88
QY      421 CTGATACACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAAATGCAATGAA 480
Db      89 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 108
QY      481 ACCATGCCCTATGTTTAAAGAAAGACGAGATTGGCGCTTCCGATGGATGAGATAGAAA 540
Db      109 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 128
QY      541 TCTACTTTTGGGAAAGAGTGGTGGCTTTCTCTGTGTAGAGAGGAGTTTCTTCTCAGCA 600
Db      129 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGluValPhePheSerGly 148
QY      601 TCTTTTCTGCTGCTATATTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCCTTTTTC 660
Db      149 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 168
QY      661 AATGAATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTGCTGATGTTCCAA 720
Db      169 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 188
QY      721 TACTTAGTAATAAGCCTTCAGAGAAAGGTCAGGAGATCATTTGTTGATGCTGCTCAAA 780
Db      189 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 208
QY      781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTTGGCCTCATTTGGAATGAATGCAATT 840
Db      209 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 228
QY      841 TTGATGAAACAGTACATGATGTTGTAGCTGACAGATPACTTGTGGAACCTTGGATCTCA 900
Db      229 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 248
QY      901 AAGGTTTTCAGGCAGAGAAATCTTTTGAATTTTATGAAACATTTCTTTTGAAGAGAAA 960
Db      249 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 268
QY      961 ACAATTTCTTTTGAGAAACAGATTTTCAGAGTATCAGCGCTTTTTCAGGTTATGCGAGAAACC 1020
Db      269 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 288
QY      1021 ACAGATAAACCTCTTTCACCTTTGGATGCGAGATTTT 1053
Db      289 ThrAspAsnValPheThrLeuAspAlaAspPhe 299

```

RESULT 6

```

Q86YB3 PRELIMINARY; PRT; 285 AA.
AC Q86YB3;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)

```


thioredoxin disulfide + H(2)O = ribonucleoside diphosphate + thioredoxin.
 CC -!- COFACTOR: Binds 2 iron ions per subunit (By similarity).
 CC -!- PATHWAY: DNA replication pathway; first step.
 CC -!- SUBUNIT: Heterodimer of a large (M1) and a small chain (M2).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- DISEASE: Ribonucleotide reductase is thought to mediate the
 CC pathogenesis of the immunodeficiency of adenosine deaminase or
 CC purine nucleoside phosphorylase. The deoxynucleotides that
 CC accumulate in the lymphoid cells of these patients are thought to
 CC feed-back inhibit ribonucleotide reductase, preventing DNA
 CC replication and cell proliferation.
 CC -!- MISCELLANEOUS: Two distinct regulatory sites have been defined:
 CC the specificity site, which controls substrate specificity, and
 CC the activity site which regulates overall catalytic activity. A
 CC substrate-binding catalytic site, located on M1, is formed only in
 CC the presence of the second subunit M2.
 CC -!- SIMILARITY: Belongs to the ribonucleoside diphosphate reductase
 CC small chain family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

EMBL; X59618; CAA42181.1; -;
 DR EMBL; AY032750; AAK51163.1; -;
 DR EMBL; BC001886; AAH01886.1; -;
 DR EMBL; BC030154; AAH30154.1; -;
 DR PIR; S25854; S25854.
 DR HSSP; P11157; 1XSW.
 DR Genew; HGNC:10452; RRM2.
 DR H-InvDB; HIX0001816; -;
 DR Reactome; P31350; -;
 DR MIM; 180390; -;
 DR GO; GO:0005737; Cytoplasm; NAS.
 DR GO; GO:0004748; Eribonucleoside-diphosphate reductase activity; NAS.
 DR GO; GO:006260; P:DNA replication; NAS.
 DR InterPro; IPR009078; Ferritin/RR like.
 DR InterPro; IPR000358; Ribonucleotide reductase.
 DR Pfam; PF00268; Ribonucleotide reductase; 1.
 DR PROSITE; PS00368; RIBRED_SMALL; 1.
 KW DNA replication; Iron; Metal-binding; Oxidoreductase.
 FT METAL 138 138 Iron 1 (By similarity).
 FT METAL 169 169 Iron 1 and 2 (By similarity).
 FT METAL 172 172 Iron 1 (By similarity).
 FT METAL 222 222 Iron 2 (By similarity).
 FT METAL 266 266 Iron 2 (By similarity).
 FT METAL 269 269 Iron 2 (By similarity).
 FT ACT SITE 176 176 By similarity.
 SQ SEQUENCE 389 AA; 44877 MW; 1056F5F84D34DA94 CRC64;

Alignment Scores:
 Pred. No.: 5,92e-119 Length: 389
 Score: 1454.00 Matches: 269
 Percent Similarity: 94.08% Conservative: 33
 Best Local Similarity: 83.80% Mismatches: 19
 Query Match: 78.76% Indels: 0
 DB: 1 Gaps: 0

US-10-698-228-12 (1-1053) x RIR2_HUMAN (1-389)
 QY 91 GAAGAGCCACTCTCAAGAAAGAGTTCTCGCGGTTTGTCTATCTTCCATTCAGTACCT 150
 Db 69 AspGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGluTrpHis 88
 QY 151 GATATTTGAAATGTATAACAGGACAGCGGTTCTCTCGACACAGAGAGGTTGAC 210
 Db 89 AspIleTrpGlnMetTyrIleLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 108

QY 211 TTATCAAGGATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAAGTACTTCTCTCT 270
 Db 109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128
 QY 271 CACATCTTAGCCTTTTGGACGACGATGGAATTTGTAATGAAAAATTTGGTGGAGCGC 330
 Db 129 HisValLeuAlaPhePheAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148
 QY 331 TTTAGTCAGAGGTCAGGTTCCAGAGCTCGCTGTTTCTATGGCTTTCAAAATTCATC 390
 Db 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
 QY 391 GAGAAATCTTCACTCAGAGATGTACAGTTTGTGTAGACACTTACATCAGAGATCCCAAG 450
 Db 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
 QY 451 AAAAGGAAATTTTATTAATGCAATTAACCAATGCAATGCTATGTTTAAAGAAAAAGAGAT 510
 Db 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysLysAlaAsp 208
 QY 511 TGGCCCTTGGATGGATAGCAGATAGAAATCTACTTTTGGGAAAGAGTGGTGGCTTT 570
 Db 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValValAlaPhe 228
 QY 571 GCTGCTCTAGAGAGTCTTCTCTCAGGATCTTTTGTCTGTATATTCTGGCTAAAGAG 630
 Db 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLys 248
 QY 631 AGAGGCTTATGCGCAGGACTCACATTTTCCAAATGAACATCATCAGCAGAGATGAAGACAT 690
 Db 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
 QY 691 CACTGTGACTTGTCTGCTGATGTTTCAATATCTAGTAATAGCCTTCAGAGAAAGG 750
 Db 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
 QY 751 GTCAGGAGATCATTTGTTGATGCTGCTCAAAATTCAGCAGGAGTCTTTTAAACAGAGCCTTG 810
 Db 289 ValArgGluIleIleIleAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 308
 QY 811 CCAGTTTGGCCTCATTTGGAATGAATTTGATTTGATGAACAGTACATTTGAGTTTGTAGCT 870
 Db 309 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 328
 QY 871 GACAGATCTTGTGGAACTTGAATTCCTCAAGGTTTTCAGGAGAGAAAAATCTCTTGTAT 930
 Db 329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348
 QY 931 TTTATGGAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACGAGTTTCAGAG 990
 Db 349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368
 QY 991 TATCAGGCTTTTTCAGTTTATGCGGAAACCAACAGATCAACGCTTTCACCTTGGATCGAGAT 1050
 Db 369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
 QY 1051 TTT 1053
 Db 389 Phe 389
 RESULT 8
 RIR2_MOUSE
 ID RIR2_MOUSE STANDARD; PRT; 390 AA.
 AC P11157;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Ribonucleoside-diphosphate reductase M2 chain (EC 1.17.4.1)
 DE (Ribonucleoside reductase small chain).
 GN Name=Rrm2;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.


```
Db 130 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 149
Qy 331 TTTAGTCAGGAGGTCAGAGGTCCTGTTCTTCTATGGCTTCAAAATTTCTCATC 390
Db 150 PheSerGlnGluValGlnValThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 169
Qy 391 GAGAAATGTTCACTCAGAGATGTACAGTGTGCTGTAGACACTTACATCAGAGATCCCAAG 450
Db 170 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 189
Qy 451 AAAAGGAATTTTATTAATGCAATTAATGAACATGCTTATGTTAAAGAAAAAGCAGAT 510
Db 190 GluArgGluTyrLeuPheAsnAlaIleGluThrMetProCysValLysLysLysAlaAsp 209
Qy 511 TGGGCTTGGAGTATGACAGATGACAGATGAAATCTACTTTGGGAAAGATGCTGGCTTT 570
Db 210 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValValAlaPhe 229
Qy 571 GCTGCTGTAGAGGAGTTTCTTCTCAGGATCTTTTGTGCTATATCTTGGCTAAAGAG 630
Db 230 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 249
Qy 631 AGAGTCTTATGCCAGGACTCACTTTTCCATGAATCACTCAGCAGAGATGAAGCACTT 690
Db 250 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 269
Qy 691 CACTGTGACTTGTGCTGCTGATGTTCCTCAATCTACTTAGTAATAAGCTTCAAGAAAGG 750
Db 270 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProAlaGluGlnArg 289
Qy 751 GTCAGGGAGATCATGTTGATGCTCTCAAAATGAGCAGGAGTCTTAAACAGAGCTTG 810
Db 290 ValArgGluIleIleThrAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 309
Qy 811 CCAGTTGGCTCATTTGGAATGAATTCATTTTGTGATGAACAGTACATTCAGTTTGTAGCT 870
Db 310 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 329
Qy 871 GACAGATTAATGTCAGTGTGATCTCAAGGTTTTCAGCGCAGAAATCTTTTGTAT 930
Db 330 AspArgLeuMetLeuGluGlyPheAsnLysIlePheArgValGluAsnProPheAsp 349
Qy 931 TTTATGGAACATTTCTTTAGAGGAAAAACAAATTTCTTTGAGAAACGAGTTTCAGAG 990
Db 350 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 369
Qy 991 TATCAGCGTTTGGCAGTTATGCGAGAAACACAGATAAGCTTTCACCTTGATGCAGAT 1050
Db 370 TyrGlnArgMetGlyValMetSerAsnSerThrGluAsnSerPheThrLeuAspAlaAsp 389
Qy 1051 TTT 1053
Db 390 Phe 390

RESULT 9
Q68EP1 ID Q68EP1 PRELIMINARY; PRT; 386 AA.
AC Q68EP1;
DT 25-OCT-2004 (TREMELrel. 28, Created)
DT 25-OCT-2004 (TREMELrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMELrel. 28, Last annotation update)
DE Hypothetical protein.
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_taxid=8364;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
```

```
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RP [2]
RC SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Gerhard D.S.;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC080161; AAH80161.1; -
DR InterPro; IPR009078; Ferritin/RR like.
DR InterPro; IPR000358; Ribonuc_redctse.
DR Pfam; PF00268; Ribonuc_red_sm; 1.
DR PROSITE; PS00368; RIBORED_SMALL; 1.
KW Hypothetical protein.
SQ SEQUENCE 386 AA; 44540 MW; 277DF859B6053824 CRC64;
```

```
Alignment Scores:
Pred. No.: 1,23e-117 Length: 386
Score: 1439.00 Matches: 268
Percent Similarity: 89.50% Conservative: 39
Best Local Similarity: 78.13% Mismatches: 36
Query Match: 77.95% Indels: 0
DB: 2 Gaps: 0
```

US-10-698-228-12 (1-1053) x Q68EP1 (1-386)

```
Qy 25 GGGCGCGGCTGGATCAGGATGAGATCATCTTCACACACCAAGGAAAGTGAATAAG 84
Db 44 AlaArgAsnIlePheGlnGluAlaGluThrTyrLysSerLysAlaProLysAspProArg 63
Qy 85 TCAATGAAAGAGACCTCTCTAAGAAAGAGTCTCGCGGTTTGTTCATCTTTCCAAATCCAG 144
Db 64 IleGlnAspGluProLeuLeuLysAspAsnProHisArgPheValIlePheProIleGln 83
Qy 145 TACCTCGATATTGGAAAAATGTATAACAGCGACAGCTTCTCTTCTGGACAGAGAGAG 204
Db 84 TyrHisAspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGlu 103
Qy 205 GTTGACTTATCAAGGATCTCCCTCACTCGAACACAGCTTAAGCAGATGAGNACTATTC 264
Db 104 ValAspLeuSerLysAspLeuArgHisTrpGluSerLeuLysAlaGluGluLysTyrPhe 123
Qy 265 ATCTCTCATCATCTTAGCTTTTGGAGCCAGGTGATGGAATTTGTAATGAAATTTGGTG 324
Db 124 IleSerHisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuVal 143
Qy 325 GAGCGCTTTAGTCAGGAGGTGCGAGGTCCAGAGCGTCTGCTTCTATGCTTTCAAAAT 384
Db 144 GluArgPheSerLysGluValGlnValThrGluAlaArgCysPheTyrGlyPheGlnIle 163
Qy 385 CTCATCGAGATGTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAT 444
Db 164 AlaMetGluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAsp 183
Qy 445 CCCAAGAAAAGGAAATTTTATTATTAATGCAATGAAACCATGCCCTATGTTAAAGAAAAA 504
Db 184 ProLysGluArgGluTyrLeuPheAsnAlaIleGluThrLeuProCysValLysLysLys 203
```


Db 211 IleSerAspLysGlnIaThrTyrGlyGluArgValValaPheAlaValaGluGly 230
 QY 586 GTTTCTCTCAGGATCTTTGCTGCTATATCTCGCTAAAGAGAGAGCTTATGCGCA 645
 Db 231 IlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLysArgGlyLeuMetPro 250
 QY 646 GGACTCCTTTTCCAATGAACATCATCAGCAGAGATGAAGGACTTCACTGCTTGCT 705
 Db 251 GlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAla 270
 QY 706 TGCTGATGTTCCAACTTAGTAATAACCTTCAGAGAAAGGTCAGGAGATCATTT 765
 Db 271 CysLeuMetPheArgHisLeuValHisLysProSerGluGluArgValValGlnLeuIle 290
 QY 766 GTTGATGCTGTCAAAATTGAGCAGGAGTGTTTTAAACAGAGCCCTTGCAGTTGCCCTCAT 825
 Db 291 ThrAsnAlaValGlnIleGluGlnGluPheLeuThrGluAlaLeuProValAsnLeuIle 310
 QY 826 GGAATGAATTTGCAATTTGATGAACACAGTACATTCAGTTTGTAGCTGACAGATTAATTG 885
 Db 311 GlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuLeu 330
 QY 886 GAATCTGGATCTCAAGGTTTTCAGTCAGCGAGAAATCCTTTGATTTATGGAACATTT 945
 Db 331 GluLeuGlyPheAsnLysValPheLysAlaSerAsnProPheAspPheMetGluAsnIle 350
 QY 946 TCTTTAGAGGAAACAAATTTCTTTGAGAAACGATTTTCAGATATCAGCGTTTGTGCA 1005
 Db 351 SerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGluTyrGlnLysMetGly 370
 QY 1006 GTTATGCGCAGAACCAACACAGATACGTCCTTCACTTCGTGATGTCAGATTTT 1053
 Db 371 ValMetSerLysProLysAspAsnThrPheThrLeuAspAlaAspPhe 386
 RESULT 11
 Q8AVY2
 ID Q8AVY2 PRELIMINARY; PRT; 386 AA.
 AC Q8AVY2;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE MGS2676 protein.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner K.H., Schenken C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Sapletenko M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Munz D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalek U., Smaluk D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.

RC TISSUE=Embryo;
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.,
 RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT initiative."
 RL Dev. Dyn. 225:384-391(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RA Klein S., Strausberg R.;
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC041209; AAH41209.1; -;
 DR HSP; P11157; IXS.
 DR GO; GO:0004748; P:ribonucleoside-diphosphate reductase activity; IEA.
 DR GO; GO:0009186; P:deoxyribonucleoside diphosphate metabolism; IEA.
 DR InterPro; IPR009078; Ferritin/RR like.
 DR InterPro; IPR000358; Ribonuclease.
 DR Pfam; PF00268; Ribonuclease sm; 1.
 DR PROSITE; PS00368; RIBOED SMALL; 1.
 SQ SEQUENCE 386 AA; 44635 MW; 4635FCC9BB18EA32 CRC64;
 Alignment Scores:
 Pred. No.: 1,4e-116 Length: 386
 Score: 1427.00 Matches: 266
 Percent Similarity: 89.88% Conservative: 36
 Best Local Similarity: 79.17% Mismatches: 30
 Query Match: 77.30% Indels: 4
 DB: 2 Gaps: 1
 US-10-698-228-12 (1-1053) x Q8AVY2 (1-386)
 QY 58 TCAGACACCAACCAAGTGAATAAAGTCAAT-----GAAGAGCCACTCTCA 105
 Db 51 ThrGluTrpLysSerLysAlaProLysAsnProArgLeuGluAspGluProLeuLeu 70
 QY 106 AGAAGAGATTCTCGCGGTTTGTTCATCTTCCATCCAGTACCTGATATTGGAAATG 165
 Db 71 LysAsnProHisArgPheValIlePheProIleGlnTyrHisAspIleTrpGlnMet 90
 QY 166 TATAACAGGCACAGGCTTCTTCTGACAGCAGCAGAGGTTTGACTTATCAAGGATCTC 225
 Db 91 TyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeu 110
 QY 226 CTTCACTGGAACCAAGCTTAAAGCAGATGAGAGTCTCTCATCTCTCATCTTACCTTT 285
 Db 111 GlnHisTrpGluSerLeuLysLysGluGluLysTyrPheIleSerHisValLeuAlaPhe 130
 QY 286 TTTGACGACGATGATGGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGCTG 345
 Db 131 PheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArgPheSerLysGluVal 150
 QY 346 CAGGTTCCAGAGCTCGCTCTTCTATGGCTTTTCAAAATTTCTCATCGAATGTTCACTCA 405
 Db 151 GlnValThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMetGluAsnIleHisSer 170
 QY 406 GAGATGTACAGTTTGTGTATAGACACTTACATCAGATGATCCCAAGAAAGGAAATTTTA 465
 Db 171 GluMetTyrSerLeuLeuIleAspThrTyrValLysAspProLysGluArgGluTyrLeu 190
 QY 466 TTTAATGCAATTGAAACCATGCTTATGTAAGAAAAGCAGATTTGGCTTGGATGG 525
 Db 191 PheAsnAlaIleGluThrLeuProCysValLysLysLysAlaAspTrpAlaLeuHisTrp 210
 QY 526 ATAGCAGATGAAATCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAAGCA 585
 Db 211 IleGlyAspLysGlnAlaThrPheGlyGluArgValValAlaPheAlaValaGluGly 230
 QY 586 GTTTCTCTCAGGATCTTTTGTGCTATATCTTGGCTAAAGAGAGAGGCTTTATGCGCA 645
 Db 231 IlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLysArgGlyLeuMetPro 250
 QY 646 GGACTCCTTTTCCAATGAACATCATCAGCAGATGAGGACTTCACTGCTTGCT 705

Db 251 GlyLeuThrPheSerAsnGluLeuLeuSerArgAspGluGlyLeuHisCysAspPheAla 270
 QY 706 TGCCTGATGTTCCAACTAGTAATAAGCCTTCAGAGAAAGGCTCAGGAGATCATT 765
 Db 271 CysLeuMetPheLysHisLeuLeuArgLysProSerGluGluValGluLeuLeu 290
 QY 766 GTTGATGCTGTCAAATTCAGACGAGGAGTTTAAACAGAGACCTTCGACGTCCTCATT 825
 Db 291 ThrAspAlaValGlnIleGluGlnGluPheLeuThrGluAlaLeuProValAsnLeuLeu 310
 QY 826 GGAATGAATTCATTTTGATGAACAGTACATTGAGTTTGTAGCTGCACAGATTACTTGTG 885
 Db 311 GlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuLeu 330
 QY 886 GAACCTGGATTCCTCAAGCTTTTTCAGGACGAGAAATCCTTTTGTATTTATGGAACATT 945
 Db 331 GluLeuGlyPheLysLysValPheLysAlaThrAsnProPheAspPheMetGluAsnIle 350
 QY 946 TCTTTAGAGGAGAAACAAATTTCTTTGAGAACGAGTTTCAGAGTATCAGCGCTTTTGCA 1005
 Db 351 SerLeuGluGlyLysThrAsnPhePheGluLysLysValGlyGluTyrGlnLysMetGly 370
 QY 1006 GTTATGGCAGAACCCACAGATAACGCTTTCACCTTCGATGCAGATTTT 1053
 Db 371 ValMetSerLysAlaLysAspAsnThrPheThrLeuAspAlaAspPhe 386

RESULT 12
 RIR2_BRARE STANDARD; PRT; 386 AA.
 AC P79733;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DE 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Ribonucleoside-diphosphate reductase M2 chain (EC 1.17.4.1)
 DE (Ribonucleotide reductase protein R2 class I).
 GN Name=r2m2; Synonyms=r2;
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97137859; PubMed=8983196;
 RA Mathews C.Z., Sjoberg B.-M., Karlsson M.;
 RT "Cloning and sequencing of cDNAs encoding ribonucleotide reductase
 RT from zebrafish Danio rerio.";
 RL Mol. Mar. Biol. Biotechnol. 5:284-287(1996).
 CC -!- FUNCTION: Provides the precursors necessary for DNA synthesis.
 CC -!- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate +
 CC thioredoxin disulfide + H(2)O = ribonucleoside diphosphate +
 CC thioredoxin.
 CC -!- COFACTOR: Binds 2 iron ions per subunit (By similarity).
 CC -!- PATHWAY: DNA replication pathway; first step.
 CC -!- SUBUNIT: Heterodimer of a large (M1) and a small chain (M2).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: Belongs to the ribonucleoside diphosphate reductase
 CC small chain family.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U57965; AAB37103.1; -.
 CC HSSP; P11157; 1XSM.
 DR ZFIN; ZDB-GENE-990415-25; r2m2.
 DR InterPro; IPR009078; Ferritin/RR like.
 DR InterPro; IPR000358; Ribonucl_reductse.

DR Pfam; PF00368; Ribonuc red sm; 1.
 KW PROSITE; PS00368; RIBRED SMALL; 1.
 FT DNA replication; Iron; Metal-binding; Oxidoreductase.
 FT METAL 135 135 Iron 1 (By similarity).
 FT METAL 166 166 Iron 1 and 2 (By similarity).
 FT METAL 169 169 Iron 1 (By similarity).
 FT METAL 229 229 Iron 2 (By similarity).
 FT METAL 263 263 Iron 2 (By similarity).
 FT METAL 266 266 Iron 2 (By similarity).
 FT ACT_SITE 173 173 By similarity.
 SQ SEQUENCE 386 AA; 44594 MW; C16846FB57F9F4E CRC64;

Alignment Scores:
 Pred. No.: 1.89e-116 Length: 386
 Score: 1425.50 Matches: 268
 Percent Similarity: 89.41% Conservative: 36
 Best Local Similarity: 78.82% Mismatches: 31
 Query Match: 77.22% Indels: 5
 DB: 1 Gaps: 1

US-10-698-228-12 (1-1053) x RIR2_BRARE (1-386)

QY 49 AGATCATCTTCAGACACCAACGAAAGTGAATAAAGTCA-----AATGAA 93
 Db 47 ArgLysIlePheAspGluSerGluGlyGlnSerLysAlaLysLysGlyAlaValGluGlu 66
 QY 94 GAGCCACTCTAAGAAAGAGATTCTCGCCGGTGTTCATCTTCCAAATCCAGTACCTCAT 153
 Db 67 GluProLeuLeuLysGluAsnProHisArgPheValIlePheProIleGlnTyrHisAsp 86
 QY 154 ATTTGAAAATGTATAAACAGGACAGGCTTCCTCTCGGACAGCAGAGAGGTTTGACTTA 213
 Db 87 IleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluValAspLeu 106
 QY 214 TCAAGAGATCTCCCTCACTCGAACAAGCTTAAAGACAGATGAGAGTACTTCATCTCAC 273
 Db 107 SerLysAspLeuGlnHisTrpAspSerLeuLysAspGluArgTyrPheIleSerHis 126
 QY 274 AICTTAGCCTTTTTCGACGACGCTGAGTAATGTAATGAAATTTGTTGGAGCGCTTT 333
 Db 127 ValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArgPhe 146
 QY 334 AGTCAGGAGGTGCGAGTTCAGAGGCTCGCTGTTCTTATGGCTTCAAAATTCATCGAG 393
 Db 147 ThrGlnGluValGlnValThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMetGlu 166
 QY 394 AATGTTCACTCAGAGATGTACAGTTTGTCTGATGACACTTACATCAGAGATCCCAAGAA 453
 Db 167 AsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspSerLysGlu 186
 QY 454 AGGGAATTTTATTAATGCAATTAACCATGCCCTATGTTAAGAAAAGCAGATGG 513
 Db 187 ArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysLysAlaAspTrp 206
 QY 514 GCCTTCGAGTGGATAGCAGATGAGAAATCTACTTTTGGGAAAGAGTGGTGGCTTTGCT 573
 Db 207 AlaLeuAsnTrpIleGlyAspLysAsnAlaArgTyrGlyGluArgValAlaPheAla 226
 QY 574 GCTGTAGAAGGAGTCTTCTCTCAGGATCTTTGCTGCTATATTCTCGCTAAAGAGAGA 633
 Db 227 AlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysArg 246
 QY 634 GGTCTTATGCCAGACTCATTCTTTTCAATGAATCTACAGCAGAGATGAAGGACTTCAC 693
 Db 247 GlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGlyLeuHis 266
 QY 694 TGTGACTTGTGCTGCTGATGTTCCCAATCTAGTAATAAGCCTTCAGAGAAAGGCTC 753
 Db 267 CysAspPheAlaCysLeuMetPheLysHisLeuIleAsnLysProSerGluGluThrVal 286
 QY 754 AGGAGATCATTTGTTGATGCTGCAAAATTCAGCAGGAGTCTTTTAAACAGAGCTTCGCA 813
 Db 287 LysLysIleIleMetAsnAlaValGluIleGluGlnGluPheLeuThrAspAlaLeuPro 306

Cy	898	TCAAAGTCTTTTCAGGCAGAAAAATCCCTTTTGATTATTGAAAAACATTCTTTAGAAGGA	957
	:	:::::	
Dd	339	AsnLysIlePheLysValGIuAenProPheAspPheMetGIuAenIIeSerLeuGlugly	358
Cy	958	AAACCAAAATCTTTGTGAACAACGAGTTTCAGAGTATCAGCGTTTTCAGATTATGGCAGAA	1017
Dd	359	LysThrAsnPhePheGluLyAsrGValgyltyrGlnArgMetGlyValMetSer---	377
Cy	1018	ACCACAGATAACGCTCTTCACCTTCGATGCAGATTTT	1053
Dd	378	-----AsnSerPheThrLeuAspAlaaspPhe	386
 RESULT 14 Q6DI44 PRELIMINARY; PRT; 386 AA.			
ID	Q6DI44	PRELIMINARY;	PRT; 386 AA.
AC	Q6DI44		
DT	25-OCT-2004	(TrEMBLrel. 28, Created)	
DT	25-OCT-2004	(TrEMBLrel. 28, Last sequence update)	
DT	25-OCT-2004	(TrEMBLrel. 28, Last annotation update)	
DE	Ribonucleotide reductase M2 polypeptide.		
GN	Name=rm2;		
OS	Brachydanio rerio (Zebrafish) (Danio rerio).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;		
OC	Cyprinidae; Danio.		
OX	NCBI_TaxID=7955;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RS	TISSUE=Embryo;		
RX	MEDLINE=223386257; PubMed=12477932; DOI=10.1073/pnas.2426038999;		
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,		
RA	Klausner R.D., Colling P.S., Wagner L., Shenmen C.M., Schuler G.D.,		
RA	Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,		
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heise F.,		
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,		
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,		
RA	Brownstein M.J., Ustin T.B., Toshilyuki S., Carninci P., Prange C.,		
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,		
RA	Boak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,		
RA	Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,		
RA	Fahay J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,		
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,		
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,		
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,		
RA	Krzywinski M.I., Skalska U., Smalls D.E., Schnrerch A., Schein J.E.,		
RA	Jones S.J., Marra M.A.;		
RT	"Generation and initial analysis of more than 15,000 full-length human		
RL	and mouse cDNA sequences."		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RS	TISSUE=Embryo;		
RC	TISSUE=Embryo;		
RL	Strausberg R.;		
RL	Submitted (Jul-2004) to the EMBL/GenBank/DDBJ databases.		
DR	EMBL; BC075746; AAH75746.1; -.		
DR	GO; GO:0004748; P:ribonucleoside-diphosphate reductase activity; IEA.		
DR	GO; GO:0009146; P:deoxyribonucleoside diphosphate metabolism; IEA.		
DR	InterPro; IPR009078; Ferritin/RR like.		
DR	InterPro; IPR000358; Ribonucl_redctase.		
DR	Pfam; PF00268; Ribonuc_red_sm; 1.		
DR	PROSITE; PS00369; RIBORED_SMALL; 1.		
SQ	SEQUENCE 386 AA; 44567 MW; 2CCACD442ADD1566 CRC64;		

Db	97	AspLeuSerLysAspLeuProHisTrpGluAlaLeuLysProGluGluArgTyrPheIle	116
QY	268	TCCTCACATCTTACGCTTTTTCAGCCAGCTGATGGAATTGTAATCAAAATTTGGTGGAG	327
Db	117	SerTyrValLeuAlaPheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGlu	136
QY	328	CGCTTTAGTCAGAGGTCCAGAGGTCGCTGTTTCTATGGCTTCCAAATTC	387
Db	137	ArgPheSerGlnGluValGlnValThrGluAlaArgCysPheTyrGlyPheGlnIleAla	156
QY	388	ATCGAGAATGTCATCTCAGAGATGTCAGTTTCTGCTGATAGACACATTCATCAGATCCC	447
Db	157	MetGluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspPro	176
QY	448	AAGAAAAGGAAATTTTATTATGCAATTGAAACCATGCCCTATGCTTAAGAAAAAGCA	507
Db	177	LysGluArgGluPheLeuPheAsnAlaIleGluThrLeuProCysValLysLysAla	196
QY	508	GATTGGCCCTTCGATGGATAGCAGATAGAAAATCTACTTTGGGAAAAGAGTGGGCC	567
Db	197	GluTrpAlaLeuArgTrpIleSerAspArgGluAlaProPheCysGluArgIleValAla	216
QY	568	TTTGCTGCTGTAGAAGGATTTCTCTCAGATCTTTTGTGCTGCTATATCTGGCTAAAG	627
Db	217	PheAlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLys	236
QY	628	AAGAGAGGTCTTATGCCAGACTCACTTTTCCAACTCATCAGCAGAGATGAAGGA	687
Db	237	LysArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGly	256
QY	688	CTTCACGTGACTTTGCTTGCCTGATGTTTCCAATACTTAGTAATTAAGCTTTCAGAA	747
Db	257	LeuHisThrAspPheAlaCysLeuLeuPheLysHisValValHisLysProSerGluGln	276
QY	748	AGGTCAGGAGATCATTTGTTGATGCTGTCAAAATTGACAGGAGTTTTTAACAGAGCC	807
Db	277	ArgValArgGluIleIleIleAspAlaValLysIleGluGlnGluPheLeuThrGluAla	296
QY	808	TTGCCAGTTTGGCTCATTCGGAATGAATTCATTTTGATGAAACAGTACATTCAGTTTGA	867
Db	297	LeuProValGluLeuIleGlyMetAsnCysLysLeuMetLysGlnTyrIleGluPheVal	316
QY	868	GCTGACAGATTACTTGTGGAACCTTGGATTCCTCAAGTGTTCAGCCAGAAAAATCCTTTT	927
Db	317	AlaAspArgLeuLeuLeuGluLeuGlyPheLysLysIlePheLysSerGluAsnProPhe	336
QY	928	GATTTTATGAAAACATTTCTTTAGAGGAAAAACAATTTCTTTGAGAAACGAGTTTCA	987
Db	337	AspPheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysLysAla	356
QY	988	GAGTATCACGCTTTTGCACTTATGGCAGAAACACACAGATAACGCTCTTCCACTTGGATGCA	1047
Db	357	GluTyrGlnLysMetGlyValMetSerLysAlaThrAspLysGlnPheThrLeuAsnAla	376
QY	1048	GATTTT 1053	
Db	377	GluPhe 378	

US-10-698-228-12 (1-1053) x O6P876 (1-378)

49	QY	AGATCATCTTCAGACACACCGAAAGTGAATA	-----AAGTCA	87
37	Db	ArgThrSerArgGlyThrHisLysGlnTrpValCysGlnSerLeuLysAspProArgGlie	56	
88	QY	AATGAAGAGCCACTCTTAAGAAAGAGTCTCGCGGTTTGTTCATCTTCCAAATCCAGTAC	147	
57	Db	GlnAspGluProLeuLeuArgAspAsnProGlyArgPheValIleLeuProIleGluIyr	76	
148	QY	CCTGATATTTCGAAAAATGATAAACAGCGACAGSCTTCTTCCTTGGACAGCAGAGAAGGTT	207	
77	Db	HisAspIleTrpGlnMecTyrLysGluAlaSerPheTrpThrAlaGluGluIval	96	
208	QY	GACTTATCAAAAGGATCTCCCTCACTCGAAACAAGCTTAAAGCAGATGAGAAGTACTTTCATC	267	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 30, 2005, 06:01:36 ; Search time 15.5 Seconds
(without alignments)
10142.641 Million cell updates/sec

Title: US-10-698-228-12

Perfect score: 1846

Sequence: 1 atggcgaccggaaaggcc.....tcacctggatgcagatttt 1053

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 1027090

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10698228/runat 26102005.100609 4010/app_query.fasta_1.2446
-DB=Issued Patents AA -QFMT=fastan -SUFFIX=n2p.ra -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEA SIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10698228 @CGN 1 1 32 .runat 26102005.100609 4010 -NCPU=6 -ICPU=3
-NO MNAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA.*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1821	98.6	351	US-10-019-733-1	Sequence 1, Appli
2	1454	78.8	389	US-10-019-733-5	Sequence 5, Appli
3	1454	78.8	389	US-09-949-016-6016	Sequence 6016, Ap
4	1454	78.8	453	US-09-949-016-7896	Sequence 7896, Ap
5	1131	61.1	411	US-09-248-796A-18017	Sequence 18017, A
6	1072.5	58.1	274	US-08-307-499-19	Sequence 19, Appl
7	1072.5	58.1	274	US-08-299-268-19	Sequence 19, Appl
8	974.5	52.8	322	US-08-136-743B-2	Sequence 2, Appli
9	969.5	52.5	360	US-09-248-796A-18018	Sequence 18018, A
10	768	41.6	147	US-08-905-223-358	Sequence 358, App
11	599	32.4	149	US-09-270-767-46236	Sequence 46236, A
12	495	26.8	323	US-09-902-540-10006	Sequence 10006, A

13	381.5	20.7	445	4	US-09-328-352-4714	Sequence 4714, Ap
14	377.5	20.4	476	4	US-09-252-991A-19358	Sequence 19358, A
15	335	18.1	365	4	US-09-198-452A-1059	Sequence 1059, Ap
16	335	18.1	380	4	US-09-438-185A-986	Sequence 986, App
c 17	299.5	16.3	121	1	US-08-307-499-7	Sequence 7, Appli
c 18	299.5	16.3	121	1	US-09-299-268-7	Sequence 7, Appli
19	298	16.1	314	3	US-09-080-044-3	Sequence 3, Appli
20	298	16.1	314	3	US-09-531-857A-3	Sequence 3, Appli
21	294	15.9	56	3	US-08-905-223-362	Sequence 362, App
22	220	11.9	382	4	US-09-543-681A-4239	Sequence 4239, Ap
23	203.5	11.0	412	4	US-09-489-039A-8180	Sequence 8180, Ap
24	201.5	10.9	401	4	US-09-540-236-3566	Sequence 3566, Ap
25	200	10.8	325	4	US-09-107-532A-4443	Sequence 4443, Ap
26	195.5	10.6	320	3	US-09-092-437-2	Sequence 2, Appli
27	195.5	10.6	320	3	US-09-583-110-3858	Sequence 3858, Ap
28	195.5	10.6	320	4	US-09-107-433-3136	Sequence 3136, Ap
29	195	10.6	339	4	US-09-543-681A-5069	Sequence 5069, Ap
30	169.5	9.2	323	4	US-09-042-775-2	Sequence 2, Appli
31	167	9.0	343	3	US-09-134-001C-5484	Sequence 5484, Ap
32	163	8.8	334	4	US-09-134-000C-5796	Sequence 5796, Ap
33	161	8.7	311	4	US-09-710-279-1260	Sequence 1260, Ap
34	157	8.5	324	4	US-08-813-940-2	Sequence 2, Appli
35	156	8.5	391	4	US-09-489-039A-10121	Sequence 10121, A
36	153	8.3	104	4	US-09-230-485-5	Sequence 5, Appli
37	138	7.5	321	2	US-08-813-940-6	Sequence 6, Appli
38	127	6.9	52	3	US-08-917-299-21	Sequence 21, Appl
39	127	6.9	52	3	US-09-422-662-21	Sequence 21, Appl
40	127	6.9	52	4	US-09-730-763-21	Sequence 21, Appl
41	127	6.9	52	4	US-09-429-370-21	Sequence 21, Appl
42	103	5.6	555	3	US-09-173-151A-32	Sequence 32, Appl
43	100.5	5.4	382	4	US-09-134-000C-5889	Sequence 5889, Ap
44	99	5.4	233	3	US-08-993-825-2	Sequence 2, Appli
45	98.5	5.3	896	4	US-09-270-767-46130	Sequence 46130, A

ALIGNMENTS

RESULT 1
US-10-019-733-1
; Sequence 1, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 1
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-1

Alignment Scores:			
Pred. No.:	4,29e-210	Length:	351
Score:	1821.00	Matches:	351
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	98.65%	Indels:	0
DB:	4	Gaps:	0

US-10-698-228-12 (1-1053) x US-10-019-733-1 (1-351)

QY 1 ATGGGCGACCCGGAAAGCGCGGCGCTGGATCAGATGAGATCATCTTCA 60


```
Db 1 MetGlyAspProGluArgProGluAaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAAGGAAAGTGAATAAGTCAATGAAGACCACTCTTAAGAAGAGTTCTCCG 120
Db 21 AspThrAsnGluSerGluIleYsSerAsnGluGluProLeuLeuArgYsSerArg 40
QY 121 CGTTTGTGCATCTTTCCCAATCCAGTACCTCTATATTGGAAAATGTATAAACAAGGCACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrPlyMetTyrLysGlnAlaGln 60
QY 181 GCTTCTCTCTGACAGCAGAGAGTGTGACTTATCAAGAGTCTCCCTCAGTCAACAAG 240
Db 61 AlaSerPheThrTrpAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAGCAGATGAGAGTACTTCTCTCTCATCTTACCTTTTTCGACCCAGTGTAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY 301 GGAATTGTAAATGAAAATTTGGTGAGCGCTTTAGTCAGAGGTGCAGGTTCAGAGGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATGGCTTCAAAATCTCATCGAGAATGTTCTACTCAGAGATGTACAGTTTG 420
Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAATTTTATTTAATGCAATTGAA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysIleArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAAGAAAAACAGATTCGGCTTCGGATGGATAGCAGATGAAAA 540
Db 161 ThrMetProTyrValLysLysIleAspThrPalaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGGAAGAGTGTGGCTTTGCTGCTGTAGAGAGAGTTTCTTCTCAGCA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTGCTATATTCTGGCTTAAAGAAAGAGAGTCTTATGCCAGGACTCATTTTTCC 660
Db 201 SerPheAlaAlaIlePheThrPheLysLysIleArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACCTCATCAGCAGAGATCAAGACTTCACTGTGACTTGTGCTGTGCTGATGTTCCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAAATTAAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGTTGATGCTGTCAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleValAspAlaValLys 260
QY 781 ATTGACGAGGAGTATTTTAAAGAGCCTTGCAGTTCGCTCATTCGCTGATGATTCAT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTGGATCTCA 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTCAGGCAGAGAAAATCCTTTGATTTTATGGAAAACATTTCTTTAGAGGAAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAATTTCTTTGAAACAGAGTTTTCAGAGTATCAGCGTTTTCGAGTTTATGCAGAAACC 1020
Db 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATACGCTTTCACCTTGGATCGAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
RESULT 2
US-10-019-733-5
```

```
; Sequence 5, Application US/10019733
; Patent No. 6682917
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/019,733
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 5
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-019-733-5

Alignment Scores:
Pred. No.: 7,14e-166 Length: 389
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: 4 Gaps: 0

US-10-698-228-12 (1-1053) x US-10-019-733-5 (1-389)
QY 91 GAAGAGCCACTCTCTAAGAAAGATTCTCGCGGTTTGTCTCATCTTTCCAAATCCAGTACCCT 150
Db 69 AspGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGluTyrHis 88
QY 151 GATATTGGAATAATGTATAACAGGCACAGCTTCTCTGACAGCAGAGAGAGGTTGAC 210
Db 89 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 108
QY 211 TTATCAAGAGATCTCCTCCTCAGTAAAGCAAGCTTAAAGCAGATGAGAAAGTACTTCTCTCT 270
Db 109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128
QY 271 CACATCTTAGCTTTTTCAGCCAGTGTGGAATTTGAAATGAAATTTGTTGGTGGAGCGC 330
Db 129 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148
QY 331 TTATAGTCAGAGGTGCAGGTTCAGAGGCTCGCTGTTCTATGGCTTTTCAAAATTTCTCATC 390
Db 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
QY 391 GAGAAATGTTCACTCAGAGATGTACAGTTTCTCATGACACTTACATCAGAGATCCCAAG 450
Db 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
QY 451 AAAAGCGAAATTTTATTAAATGCAATTGAAACCATGCCCTTATGTTAAGAAAAAGCAGAT 510
Db 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 208
QY 511 TGGGCTTTCGGATGGATAGCAGATAGAAATCTACTTTTGGGGAAGAGTGGTGGCTTTT 570
Db 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaPhe 228
QY 571 GCTGCTGTAGAAGAGTCTTTTCTCTCAGGATCTTTTGTCTCTATATTTCTGCTAAAGAG 630
Db 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheThrPheLysLys 248
QY 631 AGAGGCTTTATGCCAGGACTCCTTTTTCCAATGAACTCATCAGCAGAGATGAAGGACTT 690
Db 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
```

```
QY 691 CACTGTGACTTGTGCTTGGCTGATGTTCCAAATCTAGTAAATAGCCTTCAGAGAAAGG 750
Db 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
QY 751 GTCAGGAGATCATTTGATGCTGTCAAAATTCAGCAGGAGTTTAAACAGAGCCTTG 810
Db 289 ValArgGluLeileileAsnAlaValArgileGluGlnGluPheLeuThrGluAlaLeu 308
QY 811 CCAGTTGGCTCATTTGGAATGAATTCATTTGATGAAACAGTACATTCAGTTTGTAGCT 870
Db 309 ProValLysLeuLeileGlyMetLysCysThrLeuMetLysGlnTyrileGluPheValAla 328
QY 871 GACAGATTACTTGTGGAATCTGGAATCTCAAAAGTTTTCAGCAGAGAAATCCTTTTGAT 930
Db 329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348
QY 931 TTTATGGAACAACTTCTTAGAGAAACAAATTTCTTTCAGAGAAACGAGTTTCAGAG 990
Db 349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368
QY 991 TATCAGCGTTTGCAGTTATGCGAGAAACACAGATAACGCTCTTCACTTGGATGCAGAT 1050
Db 369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
QY 1051 TTT 1053
Db 389 Phe 389

RESULT 3
US-09-949-016-6016
; Sequence 6016, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6016
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6016

Alignment Scores:
Pred. No.: 7,14e-166 Length: 389
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: 4 Gaps: 0

US-10-698-228-12 (1-1053) x US-09-949-016-6016 (1-389)
QY 91 GAAGAGCCACTCTTAAGAAAGAGTTCTCGCGGTTGTCTATCTTCCAAATCCAGTACCCT 150
Db 69 AspGluProLeuLeuArgGluAsnProArgPheValilePheProileGluTyrHis 88
QY 151 GATATTGGAATGTATAACAGGACAGGCTTCTTCGACACAGACAGAGGTTGAC 210
Db 89 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 108
QY 211 TTATCAAGAGTCTCCCTCACTGGACACAGCTTAAGACGATGAGAGTACTTCATCTCT 270
Db 211 TTTATCAAGAGTCTCCCTCACTGGACACAGCTTAAGACGATGAGAGTACTTCATCTCT 270
```

```
Db 109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128
QY 271 CACATCTTACCTTTTTCAGCCAGTATGGAATTTGAAATGAAAAATTTGGTGGAGCQC 330
Db 129 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148
QY 331 TTTAGTCAGGAGGTGTCAGGCTCCAGAGCTCGCTGTTTCTTCTAGCTTTCAAATTCATC 390
Db 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
QY 391 GAGATGTTCTCAGAGATGTACAGTTGCTGTGATAGACACTTACATCAGAGATCCCAAG 450
Db 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
QY 451 AAAAGGGAATTTTATTAATGCAATTAACCAATGCAATGCAATGCAATGCAATGCAATG 510
Db 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 208
QY 511 TGGCCCTTGGCAGTAGCAGATAGAAATCTACTTTTGGGGAAGAGTGGTGGCCTTT 570
Db 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValValAlaPhe 228
QY 571 GCTGCTGTAGAGAGTTTCTTCTCAGGATCTTTTCTGCTATATCTGCTGCTGCTGCTG 630
Db 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 248
QY 631 AGAGTCTTATGCGAGGACTCCTTTTCCAAATGCAATGCAATGCAATGCAATGCAATG 690
Db 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
QY 691 CACTGTGACTTGTGCTTGGCTGATGTTCCAAATCTAGTAAATAGCCTTCAGAGAAAGG 750
Db 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
QY 751 GTCAGGAGATCATTTGATGCTGCTCAAAATTCAGCAGGAGTTTTCAGAGAAAGCCTTG 810
Db 289 ValArgGluLeileileAsnAlaValArgileGluGlnGluPheLeuThrGluAlaLeu 308
QY 811 CCAGTTGGCTCATTTGGAATGAATTCATTTGATGAAACAGTACATTCAGTTTGTAGCT 870
Db 309 ProValLysLeuLeileGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 328
QY 871 GACAGATTACTTGTGGAATCTGGAATCTCAAAAGTTTTCAGGAGAGAAATCCTTTGAT 930
Db 329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348
QY 931 TTTATGGAACAACTTCTTAGAGAAACAAATTTCTTTCAGAGAAACGAGTTTCAGAG 990
Db 349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368
QY 991 TATCAGCGTTTGCAGTTATGCGAGAAACACAGATAACGCTTTCACCTTGGATGCAGAT 1050
Db 369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
QY 1051 TTT 1053
Db 389 Phe 389

RESULT 4
US-09-949-016-7896
; Sequence 7896, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
```

```

; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7896
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7896

```

Alignment Scores:			
Pred. No.:	7, 74e-166	Length:	453
Score:	1454.00	Matches:	269
Percent Similarity:	94.08%	Conservative:	33
Best Local Similarity:	83.80%	Mismatches:	19
Query Match:	78.76%	Indels:	0
DB:	4	Gaps:	0

US-10-698-228-12 (1-1053) x US-09-949-016-7896 (1-453)

Qy	91	GAAGAGCCACCTCTTAAGAAAGAGTTCGCCGGTTGTCACTTTCCAATCCAGTACCCT	150
Db	133	AspGluProLeuLeuArgPheValIlePheProIleGluTrpHis	152
Qy	151	GATATTGGAAAAATATAAACAGCACAGGCTTCCTCTCGACAGCAGAAGGTTGAC	210
Db	153	AspIleTrpGlnMetTrpLysAlaGluAlaSerPheTrpThrAlaGluValAsp	172
Qy	211	TTATCAAAGGATCTCCCTCACTGGACAACAGCTTAAAGCAGATGAGAAGTACTTCATCTCT	270
Db	173	LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGuargTrpPheIleSer	192
Qy	271	CACATCTTAGCTTTTTTGCCAGCCAGTGCATGAAATTGTAATGAAAATTTGCTGGAGCGC	330
Db	193	HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg	212
Qy	331	TTTAGTCAGGAGGTCAGGTCAGAGGCTCGCTGTTTCTATGGCTTTCAAATTTCTCATC	390
Db	213	PheSerGlnGluValGlnIleThrGluAlaArgCysPheTrpGlyPheGlnIleAlaMet	232
Qy	391	GAGAAATGTTCACTCAGAGATGTACAGTTTGTGTATAGACACTTACATCAGAGATCCCAG	450
Db	233	GluAsnIleHisSerGluMetTrpSerLeuLeuIleAspThrTrpIleLysAspProLys	252
Qy	451	AAAAGGAAATTTTATTAAATGCAATTCAAAACCATGCCCTATGTTAAGNAAAAAGCAGAT	510
Db	253	GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysLysAlaAsp	272
Qy	511	TGGGCTTCGATGATGATAGCAGATAGAAAATCTACTTTTGGGGAAGAGTGTGGCTTT	570
Db	273	TrpAlaLeuArgTrpIleGlyAspLysGlnAlaThrTrpGlyGuargValValAlaPhe	292
Qy	571	GCTGCTGTAGAAGGATTTTCTCTCAGATCTTTTGTCTATATCTTGGCTAAAGAAAG	630
Db	293	AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys	312
Qy	631	AGAGTCTTATGCCAGGACTCACTTTTCCAATGAATCATCAGCAGAGATGAAGACTT	690
Db	313	ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu	332
Qy	691	CACGTGCACTTGTGCTGATGTTTCCAATCACTTAGTAAATAGCCTTCAGAGAAAGG	750
Db	333	HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg	352
Qy	751	GTCAAGGAGATCATGTGTGCTCAAAATTGACAGGAGTTTTTAAACAGAGCCTTG	810
Db	353	ValArgGluIleIleIleAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu	372
Qy	811	CCAGTGTGCCCTCATTTGGAAATGCAATTTTGATGAAACAGTACATTGAGTTTGTAGCT	870
Db	373	ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTrpIleGluPheValAla	392
Qy	871	GACAGATTACTGTGGAACTTGGAAATCTCAAGGTTTTTTCAGGCAGAAAATCCTTTTGAT	930

Db	393	AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp	412
Qy	931	TTTATGGAAACATTTCTTTAGAAAGGAAAAACAAATTTCTTTTCAGAAACGAGTTTCAGAG	990
Db	413	PheMetGluAsnIleSerLeuGluGlyLysThrAsnPheGluLysArgValGlyGlu	432
Qy	991	TATCAGCGTTTTGCAGTTATGGCAGAAACACACAGATAACGTCCTTCACCTTTGGATGCAGAT	1050
Db	433	TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp	452
Qy	1051	TTT 1053	
Db	453	Phe 453	
RESULT 5			
US-09-248-796A-18017			
; Sequence 18017, Application US/09248796A			
; Patent No. 6747137			
; GENERAL INFORMATION:			
; APPLICANT: Keith Weinstock et al			
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDID.			
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS			
; FILE REFERENCE: 107196.132			
; CURRENT APPLICATION NUMBER: US/09/248,796A			
; CURRENT FILING DATE: 1999-02-12			
; PRIOR APPLICATION NUMBER: US 60/074,725			
; PRIOR FILING DATE: 1998-02-13			
; PRIOR APPLICATION NUMBER: US 60/096,409			
; PRIOR FILING DATE: 1998-08-13			
; NUMBER OF SEQ ID NOS: 28208			
; SEQ ID NO 18017			
; LENGTH: 411			
; TYPE: PRT			
; ORGANISM: Candida albicans			
US-09-248-796A-18017			
Alignment Scores:			
Pred. No.:	5,82e-127	Length:	411
Score:	1131.00	Matches:	210
Percent Similarity:	78.21%	Conservative:	52
Best Local Similarity:	62.69%	Mismatches:	67
Query Match:	61.27%	Indels:	6
DB:	4	Gaps:	3
US-10-698-228-12 (1-1053) x US-09-248-796A-18017 (1-411)			
Qy	67	AACGAAGTGAATAAAGTCAATGAA-----GAGCCACTCTCTAAGAAGAGTTCTTCGC	120
Db	77	AsnArgHisLysLeuLysGlnGluGluSerAsnGluProLeuLeuThrProAspLysThr	96
Qy	121	CGGTTTCTCATTTTCCCATCCAGTACCCCTGATATTGGAATAATGTATAAACAGGCACAG	180
Db	97	ArgHisThrIleTyrProIleLysTyrProGluLeuTrpGlnPheTyrLysLysSerLeu	116
Qy	181	GCCTTCCTTCGGACAGCAGAAAGAGTTGACTTATCAAGAGATCTCCCTCACTCG--AAC	237
Db	117	AlaSerPheTrpThrAlaGluGluLeuAspLeuSerLysAspLeuAspTrpAsnAsn	136
Qy	238	AAAGCTTAACGACATGAGAGTAGTACTTCATCTTCACATCTTAGCCCTTTTTCGAGCAGT	297
Db	137	LysMetAsnGluAsnGluArgPhePheIleSerArgValLeuAlaPhePheAlaIleSer	156
Qy	298	GATCGAATTGTAATGAATAATTTGGTGGACGCTTTAGTCAGAGCGTGCAGGTTCCAGAG	357
Db	157	AspGlyIleValAsnGluAsnLeuValGluAsnPheCysAlaGluValGlnIleProGlu	176
Qy	358	GCTCGCTGTTCTATGGCTTTCAAATTTCTCATCGAGAATGTTTCACTCAGAGATGTACAGT	417
Db	177	AlaLysLeuValTyrLysPheGlnIleMetMetGluAsnIleHisSerGluThrTyrSer	196
Qy	418	TTGCTGATAGACACTTACATCAGACATCCCAAGAAAGGGAATTTTTTATTATGCAATT	477

```
Db 197 LeuLeuIleGluThrTyrPheLysAspProGluGluAlaAspPheLeuPheAsnAlaIle 216
QY 478 GAAACCATGCCCTATGTTAAAGAAAAAGACAGATTGGCGCTTCGATGGATAGCAGATAGA 537
Db 217 AspAsnIleProPheIleArgLysAlaAspTrpAlaIleArgTrpIleGlnSerGlu 236
QY 538 AATCTACTTTGGGGAAGAGTGTGGCTTGTCTGTAGTGAAGAGTATTTCTCTCA 597
Db 237 AspAlaLeuTyrAlaGluArgLeuValAlaPheAlaAlaValGluGlyIlePhePheSer 256
QY 598 GGATCTTTTGTCTATATTCTGGCTTAAGAGAGAGAGGTCTTATGCCAGACTCATT 657
Db 257 GlySerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPhe 276
QY 658 TCAATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGGCTGATGTC 717
Db 277 SerAsnGluLeuIleCysArgAspGluGlyIleHisThrAspTrpAlaCysLeuLeuPhe 296
QY 718 CAATACTTAGTAATAAGCTTTCAGAAGAAAGGTCAGGAGATCATTTGTTGATGCTGTC 777
Db 297 SerTyrLeuLysAsnLysProSerGluIleIleGluLysIleIleThrGluAlaVal 316
QY 778 AAATTTAGCAGGAGTATTAACAGAGCCTTGCAGTGGCTCATTTGGAATGATGTC 837
Db 317 AspIleGluLysGluTyrPheSerAspAlaLeuProValSerLeuLeuGlyMetAsnCys 336
QY 838 ATTTTGTAGAAACAGTACATGAGTTGTAGCTGACAGATTAATCTGTGAACTTGGATTC 897
Db 337 AspLeuMetCysGlnTyrValGluPheValAlaAspArgLeuLeuValAlaPheGlyAsn 356
QY 898 TCAAGAGTTTTCAGGCAGAAATCCTTTTGTGATTTTATGAAAAACATTTCTTTAGAAGA 957
Db 357 LysLysTyrTyrAsnValThrAsnProPheAspPheMetGluAsnIleSerLeuAlaGly 376
QY 958 AAACAAATTTCTTTGAGAACAGCTTTTCAGAGTATCAGCGTTTTCAGATTATGCGAGAA 1017
Db 377 LysThrAsnPhePheGluLysArgValSerAspTyrGlnLysAlaGlyValMetGluLys 396
QY 1018 ACC-----ACAGATAACGTCTTCACCTTGGATGCGAGATTTT 1053
Db 397 ValGluAsnLysAsnGluLysThrGlyLeuPheAspGlnAspPhe 411

RESULT 6
US-08-307-499-19
; Sequence 19, Application US/08307499
; Patent No. 5651972
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; TITLE OF INVENTION: Live Vaccine Vector
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: U.S.A.
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/307,499
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,241
; FILING DATE: 1-JUL-1992
; PRIOR APPLICATION DATA:
```

```
; APPLICATION NUMBER: US 07/908,630
; FILING DATE: 29-JUN-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/342,212
; FILING DATE: 21-APR-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF35.1.PWCC1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 274 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-307-499-19
```

```
Alignment Scores:
Pred. No.: 5,21e-120 Length: 274
Score: 1072.50 Matches: 204
Percent Similarity: 85.56% Conservative: 27
Best Local Similarity: 75.56% Mismatches: 38
Query Match: 58.10% Indels: 1
DB: 1 Gaps: 1
```

US-10-698-228-12 (1-1053) x US-08-307-499-19 (1-274)

```
QY 94 GAGCCACTCTTAAGAAAGAGTTCTCGCGGTTTGTCTATCTTCCAAATCCAGTACCTGAT 153
Db 2 GluProIleLeuGlnGlnSerAspPheValIlePheProIleLysTyrHisAsp 21
QY 154 ATTTGAAAATGTTATAACAGGACAGGCTTCTCTGGACAGCAGCAGAGAGGTTGACTTA 213
Db 22 IleTrpLysMetTyrLysGlnSerValAlaSerPheTrpThrValGluGluValAspLeu 41
QY 214 TCAAGAGATCTCCCTCACTCGAAACAAGCTTAAAGCAGATGAGAAAGTACTTCTCTCAC 273
Db 42 SerLysAspLeuAspAspTrpAspLysLeuThrLysAspGluLysTyrPheIleLysHis 61
QY 274 ATCTTAGCCTTTTTCAGCCAGTGTAAATGTAATGAAATTTGGTGGAGCGCTTT 333
Db 62 IleLeuAlaPhePheAlaSerSerAspGlyIleValAsnGluAsnLeuAlaGluArgPhe 81
QY 334 AGTCAGGAGTGCAGGTTCCAGAGGCTTCAGAGGCTTCGCTGTTCTATGGCTTCAAAATCTCATCGAG 393
Db 82 TyrValAspValGlnCysSerGluAlaArgCysPheTyrGlyPheGlnIleAlaMetGlu 101
QY 394 AATGTTCACTCAGAGATGTACAGTTTGTCTATGACACTTACATCAGAGATCCCAAGAAA 453
Db 102 AsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrValArgAspAsnIleGlu 121
QY 454 AGGGAATTTTATTAATGCAATTGAAACCATGCCCTATGTTAAAGAAAAAGCAGATTGG 513
Db 122 LysMetHisLeuPheAsnAlaIleGluThrMetGluCysValLysLysLysAlaAspTrp 141
QY 514 GCCTTCCGATGATACAGATAGAAATCTACTTTTGGGAGAGAGTGGTGGCTTCTGCT 573
Db 142 AlaArgLysTrpIleSerSerAsnLys---ValTyrGlyGluArgValAlaPheAla 160
QY 574 GCTGTAGAAGAGTATTTCTTCTCAGGATCTTTTGTGCTATATTTCTGGCTTAAAGAGAGA 633
Db 161 AlaValGluGlyIlePhePheSerGlySerPheAlaAlaIlePheTrpIleLysLysArg 180
QY 634 GGTCTTATCCAGGACTCATTCTTTTCCAATGAATCATCAGCAGAGATGAAGGACTTCAC 693
Db 181 GlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHis 200
QY 694 TGTGACTTTGCTTGCCTGATGTTCCAAATCTACTTAGTAATAAGCCTTCCAGAGAAAGGTC 753
```

Db 201 CysAspPheAlaCysLeuMetPheLysHisLeuLeuHisProSerLysGluValIle 220
QY 754 AGGGAGATCATTTGATGCTCTCAAAATTGACGAGGATTTTAAACAGAGCTTGCCA 813
Db 221 ThrSerIleIleIleAspAlaValAsnIleGluLysGluPheLeuThrValAlaIlePro 240
QY 814 GTTGCCCTCATTTGGAATGAATTTGATGAAACAGTACATTCAGTTTGTAGCTGAC 873
Db 241 ValAspLeuIleGlyMetAsnCysCysLeuMetSerGlnTyrIleGluPheValAlaAsp 260
QY 874 AGATTACTTGTGGAATTCGATTTGATGAAACAGTACATTCAGTTTGTAGCTGAC 873
Db 261 ArgLeuLeuThrGluLeuGlyCysGluLys 270

RESULT 7

US-09-299-268-19
; Sequence 19, Application US/09299268
; Patent No. 6217882
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; TITLE OF INVENTION: Live Vaccine Vector
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSER: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: U.S.A.
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/299,268
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/901,127
; FILING DATE:
; APPLICATION NUMBER: US 07/908,241
; FILING DATE: 1-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,630
; FILING DATE: 29-JUN-1992
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/342,212
; FILING DATE: 21-APR-1992
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF35.1.FWCC1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 274 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLSCULE TYPE: protein
US-09-299-268-19

Alignment Scores:
Pred. No.: 5,21e-120 Length: 274
Score: 1072.50 Matches: 204

Percent Similarity: 85.56% Conservative: 27
Best Local Similarity: 75.56% Mismatches: 38
Query Match: 58.10% Indels: 1
DB: 3 Gaps: 1
US-10-698-228-12 (1-1053) x US-09-299-268-19 (1-274)
QY 94 GAGCCACTCTCTAAGAAAGAGTCTCGCCGGTTTGTCTCATCTTTTCCAAATCCAGTACCCCTGAT 153
Db 2 GluProIleLeuGlnGluSerAspSerArgPheValIlePheProIleLysTyrHisAsp 21
QY 154 ATTTGAAAATGTATAACAGCAGCAGGCTTCCTTCTGGACAGCAGAGAGAGGTTGACTTA 213
Db 22 IleTyrLysMetTyrLysGlnSerValAlaSerPheTyrThrValGluGluValAspLeu 41
QY 214 TCAAGAGATCTCCCTCAGTGAACAAGCTTAAGCAGATGAGAAGTACTTCTCTCAC 273
Db 42 SerLysAspLeuAspAspTyrAspLysLeuThrLysAspGluLysTyrPheIleLysHis 61
QY 274 ATCTTAGCCCTTTTTCAGCCAGTGTAGTGAATTCATAATGAAATTTGGTGGAGCGCTTT 333
Db 62 IleLeuAlaPhePheAlaSerSerAspGlyIleValAsnGluAsnLeuAlaGluArgPhe 81
QY 334 AGTCAGGAGGTGCAGGTTCCAGAGGCTCGCTGTTTCTATGGCTTTTCAAAATTTCTCATCGAG 393
Db 82 TyrValAspValGlnCysSerGluAlaArgCysPheTyrGlyPheGlnIleAlaMetGlu 101
QY 394 AATGTTCACTCAGAGATGTACAGTTCTGCTATAGACACTTACATCAGAGATCCCAAGAAA 453
Db 102 AsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrValArgAspAsnIleGlu 121
QY 454 AGGGAATTTTATTATGCAATTGAAACCATGCCCTATGTTAAGAAAAAGCAGAGATTGG 513
Db 122 LysMetHisLeuPheAsnAlaIleGluThrMetGluCysValLysLysLysAlaAspTyr 141
QY 514 GCCTTGGCATGTAGACATAGAAAATCTACTTTTGGGGAAGAGTGGTGGCCCTTTGCT 573
Db 142 AlaArgLysTyrPheSerAsnLys--ValTyrGlyGluArgValValAlaPheAla 160
QY 574 GCTGTAGAGGAGTGTCTTCTCAGGATCTTTGCTGCTATATTTCTGGCTAAAGAGAGA 633
Db 161 AlaValGluGlyIlePhePheSerGlySerPheAlaAlaIlePheTyrIleLysLysArg 180
QY 634 GGTCTTATGCCAGGACTCAGCTTTTCCAATGAACTCATCAGCAGATGAAGGACTTCAC 693
Db 181 GlyLeuMetProGlyLeuThrPheSerAsnGluLeuSerArgAspGluGlyLeuHis 200
QY 694 TGTGACTTTGCTTGCCTGTGTTTCCAATGAACTCATCAGCAGATGAAGGACTTC 753
Db 201 CysAspPheAlaCysLeuMetPheLysHisLeuHisProProSerLysGluValIle 220
QY 754 AGGGAGATCATTTGATGCTGTCAAAATTGACGAGGAGGATTTTAAACAGAGCCTTGCCA 813
Db 221 ThrSerIleIleIleAspAlaValAsnIleGluLysGluPheLeuThrValAlaIlePro 240
QY 814 GTTGGCCTCATTTGGAATGAATTCATTTTTCATGAAACAGTACATTCAGTTTGTAGCTGAC 873
Db 241 ValAspLeuIleGlyMetAsnCysCysLeuMetSerGlnTyrIleGluPheValAlaAsp 260
QY 874 AGATTACTTGTGGAATTCGATTTGATGAAACAGTACATTCAGTTTGTAGCTGAC 903
Db 261 ArgLeuLeuThrGluLeuGlyCysGluLys 270

RESULT 8

US-08-136-743B-2
; Sequence 2, Application US/08136743B
; Patent No. 5459063
; GENERAL INFORMATION:
; APPLICANT: Barry S. Cooperman, Harvey Rubin,
; APPLICANT: Jerome Salem, and Alison L. Fisher
; TITLE OF INVENTION: "Plasmodium falciparum Ribonu-
; TITLE OF INVENTION: cleotide Reductase, DNA Sequences Therefor and Peptide Inhibi-
; TITLE OF INVENTION: Thereof"

NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: The University of Pennsylvania
STREET: Suite 330
STREET: 3700 Market Street
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: U.S.A.
ZIP: 19104-3246
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/136,743B
FILING DATE: 10/14/93
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Monaco, Daniel A.
REGISTRATION NUMBER: 30,480
REFERENCE/DOCKET NUMBER: 3957-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-8383
TELEFAX: (215) 568-5549
TELEX: No. 5459063e
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 322 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-136-743B-2

Alignment Scores:			
Pred. No.:	3.6e-108	Length:	322
Score:	94.50	Matches:	191
Percent Similarity:	75.62%	Conservative:	51
Best Local Similarity:	59.6%	Mismatches:	75
Query Match:	52.7%	Indels:	3
DB:	1	Gaps:	2

US-10-698-228-12 (1-1053) x US-08-136-743B-2 (1-322)

100	QY	CTCCTTAAGAAAGAGTCTTCGCGGGTTTCATCTTCCAATCCAGTACCTCGATATTGG	159
		::: ::: ::: ::: ::: ::: ::: ::: :::	
4	Db	IleLeuAsnLysGluSerAspArgPheThrLeuTyrProIleLeuTyrProAspValPhe	23
		::: ::: ::: ::: ::: ::: ::: :::	
160	QY	AAATGCTTAACACGACGCTCCTTCGGACAGCAGACAGAGGTGTGACTTATCAAG	219
		::: ::: ::: ::: ::: ::: ::: :::	
24	Db	ProPheTyrLysLysAlaGluAlaSerPheTrpThrAlaGluLysIleAspTyrSerSer	43
		::: ::: ::: ::: ::: ::: ::: :::	
220	QY	GATCTCCCTCACTCGAACAAGCTTAAAGCAGCATCAGAAAGTACTTTCATCTCTCACATCTTA	279
		::: ::: ::: ::: ::: ::: ::: :::	
44	Db	AspLeuLysAspPheGluLysLeuAsnGluLysHisPheIleLysHisValLeu	63
		::: ::: ::: ::: ::: ::: ::: :::	
280	QY	GCCTTTTTCGACCCAGTCAGTAATGTAAATGAAAATTTGGTGAGCCCTTTAGTCAG	339
		::: ::: ::: ::: ::: ::: ::: :::	
64	Db	AlaPhePheAlaIleAspGlyIleValLeuGluAsnLeuAlaValSerPheLeuArg	83
		::: ::: ::: ::: ::: ::: ::: :::	
340	QY	GAGGTGCAGGTTCCAGAGCTCGCTGTTTTCATGGCTTTCAAATCTCATCGAAGATGT	399
		::: ::: ::: ::: ::: ::: ::: :::	
84	Db	GluValGlnIleThrGluAlaLysPheTyrSerPheGlnIleAlaValGluAsnIle	103
		::: ::: ::: ::: ::: ::: ::: :::	
400	QY	CACCTCAGAGATGTACAGTTTGCATGACACTTTCATCATCAGATCCCAAGAAAGGGAA	459
		::: ::: ::: ::: ::: ::: ::: :::	
104	Db	HisSerGluThrTyrSerLeuLeuIleAspAsnTyrIleLysAspGluLysGluArgLeu	123
		::: ::: ::: ::: ::: ::: ::: :::	
460	QY	TTTTTATTATTAATGCAATTCACCACTGTCCTTATCTTAAGAAAAAGCAGATTCGGCCTTG	519
		::: ::: ::: ::: ::: ::: ::: :::	
124	Db	AsnLeuPheHisAlaIleGluAsnIleProAlaValLysAsnLysAlaLeuTrpAlaAla	143
		::: ::: ::: ::: ::: ::: ::: :::	
520	QY	CGATGGATAGCAGATAGAAAAATCTACTTTTGGGGAAGAGTGTGGCCTTTGCTGCTGTA	579
		::: ::: ::: ::: ::: ::: ::: :::	

Db	144	LyseTrpIleAsnAspThrAsnSer---PheAlaGluArgIleValAlaAsnAlaCysVal	162
Qy	580	GAAGGAGCTTTCTTCTCAGGATCTTTTGTGCTGCTATATTCTGGCTAAAGAGAGAGAGCTCTT	639
Db	163	GlulGlyIleLysPheSerGlySerPheCysAlaIlePheTrpPheLysLysGlnAsnLys	182
Qy	640	ATGCCAGGACTCACTTTTCCAACTGAACATCATCAGCAGAGATGAAGGACTTCACTGTGAC	699
Db	183	LeuHieGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisThrAsp	202
Qy	700	TTTGCTTGCCCTGATGTCCAAATACTTAGTAAATAAGCCTTCAGAACAAAGGGTCAGGGAG	759
Db	203	PheAsnCysLeuIleTrpSerLeuLeuAspAsnLysLeuProGluGlnMetValGlnAsn	222
Qy	760	ATCAATTGTTGATGCT-----GTCAAAATGTAGCAGGAGTTTTTATACAGAAGCCTTGCCCA	813
Db	223	IleValLysGlnAlaGlyGlyValGluValGluLysSerPheIleCysGluSerLeuPro	242
Qy	814	GTTGGCCTCAATCGAATGAATTCATTTGTAGAAACAGTACATTCAGTTGTTGAGCTGAC	873
Db	243	CysAspLeuIleGlyMetAsnSerArgLeuMetSerGlnThrIleGluPheValAlaAsp	262
Qy	874	AGATTACTTGCGAACTTGGATTCTCAAGTGTGTTTTTCAGCGCAAAAATCCTTTTGATTT	933
Db	263	ArgLeuLeuGluCysLeuGlyCysSerLysIlePheHisSerLysAsnProPheAsnTrp	282
Qy	934	ATGCAAAACATTTCTTTTAGAAGGAAACAAATTTCTTTTGAGAAACAGTTTCAGAGTAT	993
Db	283	MetAspLysIleSerLeuGlnGlyLysThrAsnPheGluLysArgValAlaAspTyr	302
Qy	994	CAGCGTTTTGCAGTTATGGCAGAAACCAACAGATAACGTCTTCACTTTGGATGCAGATTTT	1053
Db	303	GlnLysSerGlyValMetAlaGlnArgLysAspHisValPheCysLeuAsnThrGluPhe	322

RESULT 9

US-09-248-796A-18018
 ; Sequence 18018, Application US/09248796A
 ; Patent NO. 6747137
 ; GENERAL INFORMATION:
 ; APPLICANT: Keith Weinstock et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA
 ; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.132
 ; CURRENT APPLICATION NUMBER: US/09/248,796A
 ; CURRENT FILING DATE: 1999-02-12
 ; PRIOR APPLICATION NUMBER: US 60/074,725
 ; PRIOR FILING DATE: 1998-02-13
 ; PRIOR APPLICATION NUMBER: US 60/096,409
 ; PRIOR FILING DATE: 1998-08-13
 ; NUMBER OF SEQ ID NOS: 28208
 ; SEQ ID NO 18018
 ; LENGTH: 360
 ; TYPE: PR1
 ; ORGANISM: Candida albicans
 ; US-09-248-796A-18018

Alignment Scores:		
Pred. No.:	1.53e-107	360
Score:	969.50	Matches: 179
Percent Similarity:	80.56%	Conservative: 53
Best Local Similarity:	62.15%	Mismatches: 53
Query Match:	52.52%	Indels: 3
DB:	4	Gaps: 2

US-10-698-228-12 (1-1053) X US-09-248-796A-18018 (1-360)

Qy	13	GAAGGCCGGAAGCGCGGCTGGATCAGATGAGAGATCATCTTCAGACACCAACGAA	72
Db	72	GLuysAspSerLeuLysLysHisGlnAspPheLeuAlaLysHisLysValHisArg	91
Qy	73	AGTGAATAAAG-----TCAAAATGAAGAGCCATCTCTAAGAAAGAGTTTCTCGCCGGTTT	126
Db	92	HisLysLysLysGlnLeuGluAlaGluProLeuLeuValGluAsnLysArgArgTyr	111

```
QY 127 GTCATCTTTCCAAATCCAGTACCTGATATTTGGAAATGTATAAACAGGCACAGGCTTCC 186
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
112 ValMetPheProIleArgTyrHisGluIleThrAsnPhetYrIysIysAlaGluAlaSer 131
QY 187 TTCTGACAGCAGAGAGGTTGACTTATCAAGGATCTCCCTCAGTGG---AACAGCTT 243
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
132 PheTrpThrAlaGluIleAspLeuSerIysAspLeuAspAspTrpAsnAsnIysLeu 151
QY 244 AAAGCAGATGAGAGTACTCTCTCATCTTACCTTTAGCCCTTTTTCAGCCAGTGTGGA 303
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
152 AsnGluAsnGluArgTyrPheIleSerArgValLeuAlaPhePheAlaAlaSerAspGly 171
QY 304 ATTGTAAATGAAATTTGGTGAGCCCTTTAGTCAGAGGTGAGTTCACAGGCTCGC 363
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
172 IleValGlyLeuAsnLeuIleGluAsnPheSerThrGluValGlnLeuProGluAlaIys 191
QY 364 TGTCTTCTATGGCTTTCAATCTTCATCGAGATGTTTCACTCAGAGATGATACAGTTTGCTG 423
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
192 SerPheTyrGlyPheGlnIleMetMetGluAsnIleHisSerGluThrTyrSerLeuLeu 211
QY 424 ATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATGAAACC 483
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
212 IleGluThrTyrIleIysAspProGlnGluAlaAspTyrLeuPheAsnAlaIleAlaAsn 231
QY 484 ATGCCCTATGTTAAGAAAAAGCAGATTCGGCTTCGGATGGATGGATGAGTAAATCT 543
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
232 IleProCysIleGlnIysAlaAspTrpAlaIleIysTrpIleGlnAspAspGluAla 251
QY 544 ACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAGGAGTTTCTTCTCAGGATCT 603
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
252 LeuTyrGlyGluArgLeuValAlaPheAlaAlaValGluGlyIlePhePheSerGlySer 271
QY 604 TTTGCTGTATATTTCTGGCTTAAAGAGAGAGGTCTTATGCGGAGTCTCACTTTTCCAAT 663
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
272 PheAlaSerIlePheTrpLeuIysIysArgGlyLeuMetProGlyLeuThrPheSerAsn 291
QY 664 GAATCATCAGCAGAGATCAAGGATCTCACTGTGACTTTGCTGTGCTGATGTTCCATAC 723
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
292 GluLeuIleCysArgAspGluGlyLeuHisThrAspPheAlaCysLeuLeuPheAlaHis 311
QY 724 TTAGTAAATAAGCCTTCAAGAAAGGTCAGGAGATCATTTGATGCTGTCAAAAT 783
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
312 LeuGlnAsnArgProSerProGluIleValGluArgIleIleThrGluAlaValAspVal 331
QY 784 GAGCAGGAGTTTAAAGAGCCTTGGCCAGTTGGCTCATTTGGAATGATGCAATTTTG 843
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
332 GluLysGluTyrPheThrAspValLeuProValSerLeuLeuGlyLeuAsnCysAsnLeu 351
QY 844 ATGAACAGTACATTCAGTTTGTGA 867
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
352 MetCysProTyrIleGluPheLeu 359
```

RESULT 10

```
US-08-905-223-358
; Sequence 358, Application US/08905223
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste D.
; APPLICANT: Duclert, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: 5' ESTs FOR SECRETED PROTEINS
; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Knobbe, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/905,223
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Israel, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 358:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 147 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; TISSUE TYPE: Brain
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: -101...-1
; IDENTIFICATION METHOD: Von Heijne matrix
; OTHER INFORMATION: score 4.4
; OTHER INFORMATION: seq ISHILAFFAASDG/IV
; US-08-905-223-358
```

Alignment Scores:

Pred. No.:	1,75e-83	Length:	147
Score:	768.00	Matches:	147
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	41.60%	Indels:	0
DB:	3	Gaps:	0

US-10-698-228-12 (1-1053) x US-08-905-223-358 (1-147)

QY	1	ATGGGCGACCGGAAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA	60
DB	1	MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer	20
QY	61	GACACCAACGAAGTGAATAAAGTCAATGAAGAGCCACTCCTAAGAAAGATTCTCGC	120
DB	21	AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg	40
QY	121	CGGTTTGTCACTTTCATCCAGTACCCCTGATATTTGGAAATGTATAAACAGGCACAG	180
DB	41	ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln	60
QY	181	GCTTCTCTTCGACAGCAGAGAGGTTGACTTATCAAAAGGATCTCCCTCAGTGGAAACAG	240
DB	61	AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys	80
QY	241	CTTAAAGCAGATCAGAAAGTACTTCATCTCTCACATCTTAGCCCTTTTTCGAGCCAGTAT	300
DB	81	LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp	100
QY	301	GGAATTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGCT	360
DB	101	GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla	120
QY	361	CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGATGTTCTCACTCAGAGATGTACAGTTG	420
DB	121	ArgCysPheTyrGlyPheGlnIleLeuIleLeuIleGluAsnValHisSerGluMetTyrSerLeu	140
QY	421	CTCATGACACTTACATCAGA 441	
DB	141	LeuIleAspThrTyrIleArg 147	

RESULT 11

```
US-09-270-767-46236
; Sequence 46236, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; PRIOR FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 46236
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-46236

Alignment Scores:
Pred. No.: 3,986-63 Length: 149
Score: 599.00 Matches: 114
Percent Similarity: 85.91% Conservative: 14
Best Local Similarity: 76.51% Mismatches: 21
Query Match: 32.45% Indels: 0
DB: 4 Gaps: 0

US-10-698-228-12 (1-1053) x US-09-270-767-46236 (1-149)
QY 586 GTTTCCTCTCAGGATCTTTGCTGCTATATCTGCTAAAGAGAGAGGTCCTATGCGCA 645
Dy 1 IIPhePheSerGlySerPheAlaSerIlePheTrpLeuLysLysArgGlyLeuMetPro 20
QY 646 GGACTCACCTTTTCAATCAATCATCAGCAGAGAGGAGGAGGCTTCACTGTGACTTTGCT 705
Dy 21 GlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAla 40
QY 706 TGCCTGATCTTCAATCTAGTAAATAAGCCTTCAGAGAGAGGCTCAGGAGAGATCAT 765
Dy 41 ValLeuMetPheGlnHisLeuValGlnArgProLysArgGluArgIleLeuGluIle 60
QY 766 GTTGATGCTGCAAAATGAGCAGGAGGTTTAAACAGAGCCTTGCCAGTGGCCTCAT 825
Dy 61 ArgAspAlaValAlaIleGluGlnPheLeuThrAspAlaLeuProValAsnLeuIle 80
QY 826 GGAATGAATTCATTTGATGAACAGTACATGATGATGATGATGATGATGATGATG 885
Dy 81 GlyMetAsnCysAspLeuMetSerGlnTyrIleGluPheValAlaAspArgLeuLeuVal 100
QY 886 GAATCTGGATCTCAAGGTTTTCAGGCAGAGAAATCCCTTTTATGTTTATGAAACAT 945
Dy 101 GluLeuGlyValGlyLysIleTyrAsnThrLysAsnProPheAsnPheMetGluMetIle 120
QY 946 TCTTTAGAGGAGAAACAAATTTCTTTGAGAACCGAGTTTCAGAGTATCAGCGTTTTC 1005
Dy 121 SerLeuAspGlyLysThrAsnPhePheGluLysValGlyGluTyrGlnArgMetGly 140
QY 1006 GTTATGGCAGAACACAGATACGTC 1032
Dy 141 ValValSerAsnProLeuAspAsnVal 149

RESULT 12
US-09-902-540-10006
; Sequence 10006, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
```

```
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 10006
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-10006

Alignment Scores:
Pred. No.: 2,01e-50 Length: 323
Score: 495.00 Matches: 113
Percent Similarity: 54.49% Conservative: 63
Best Local Similarity: 34.98% Mismatches: 125
Query Match: 26.81% Indels: 22
DB: 4 Gaps: 6

US-10-698-228-12 (1-1053) x US-09-902-540-10006 (1-323)
QY 145 TACCCTGATATTTGGAAAAATGTATAAACAGCAGCAGCGCTTCTTCTGACAGCAGAGAG 204
Dy 3 TyrProGlnPhePheGluMetTyrArgAsnAlaIleLysAsnThrTrpThrValGluGlu 22
QY 205 GTTGACTTATCAAGGATCTCCTCTACTGG---AACAGCTTAAAGCAGATGAGAGATAC 261
Dy 23 ValAspPheSerThrAspLeuValAspLeuArgSerLysMetThrAspAlaGluArgHis 42
QY 262 TTCATCTCTCATCTTAGCCTTTTTCGACCCAGTCATGGAATGTAAATGAAATTTG 321
Dy 43 LeuIleHisArgLeuValAlaPhePheAlaThrGlyAspSerIleValGlyAsnAsnLeu 62
QY 322 GTCGAGCGCTTATAGTCAGGAGGTGCAGGTTCCAGAGCTCGCTGTTCTATGCTTTC 381
Dy 63 ValLeuAsnLeuTyrLysHisLeuAsnAlaProGluAlaArgMetTyrLeuSerArgGln 82
QY 382 ATTCTCATCGAGAATGTTCATCAGAGATGTACAGTTTGTGTATAGACACATTCATCAGA 441
Dy 83 LeuTyrGluGluAlaLeuHisValGlnPheTyrLeuThrLeuLeuAspThrTyrValPro 102
QY 442 GATCCCAAGAAAAGGAAATTTTATTAATGCAATTTGAAACCATGCTTATGTTAAGAAA 501
Dy 103 AspProAlaGluArgAlaMetAlaPheAlaValAspAsnIleProSerIleGlnArg 122
QY 502 AAAGCAGATTTGGCGCTTGCATGATGATA-----GCAGATAGAAAATCTACT 546
Dy 123 LysAlaGlnPheCysMetLysTrpMetAspSerIleGlnGlyLeuAspThrLeuGlnThr 142
QY 547 TTTGGGGAAGA-----GTGTGGCTTTTGTCT---GCTGTAGAAAGGA 585
Dy 143 LysAlaHisArgArgGlnPheLeuLeuAsnLeuIleCysPheAlaGlyCysIleGluGly 162
QY 586 GTTTTCTTCTCAGGATCTTTTGTGCTATATCTGGCTAAGAAAGAGAGGCTTATGCGCA 645
Dy 163 LeuPhePhePheAlaAlaPheAlaTyrValTyrPheLeuArgSerLysGlyLeuLeuAsn 182
QY 646 GGACTCACCTTTTCCAAATGATCATCAGCAGAGATGAAGACTTCTACTGTGACTTTCCT 705
Dy 183 GlyLeuAlaAlaGlyThrAsnTrpValPheArgAspGluSerAlaHisMetAlaPheAla 202
QY 706 TGCCTGATGTTCCAATAC-----TTAGTAAATAAGCCTTCAGAA 744
Dy 203 PheGluSerIleGlnValAlaArgLysGluGluProAspLeuPheAspAlaGlnMetGlu 222
QY 745 GAAAGGTCAGGAGATCATTTGTTGATGCTGTCTCAAAATTTGAGGAGAGTCTTAAACAGA 804
Dy 223 ArgAspValValAlaMetLeuArgGluAlaValGluCysGluThrGlnPheAlaGlnAsp 242
QY 805 GCCTTGCAGTTCGCTCATTCATGAATGAATGCAATTTGATGAACACAGTACATGAGTTT 864
Dy 243 LeuLeuSerGlyGlyValMetGlyLeuSerValGlnAspMetArgGlyTyrLeuGluTyr 262
QY 865 GTAGCTGACAGATTTACTTGTGGAACCTTGGATTTCTCAAGAGTCTTTCAGGAGAGAAATCCT 924
Dy 263 ValAlaAspGlnArgLeuGlnMetLeuGlyMetThrProValPheGlyThrLysAsnPro 282
```


Alignment Scores:	3.49e-36	Length:	476
Pred. No.:	377.50	Matches:	100
Score:	45.87%	Conservative:	72
Percent Similarity:	26.67%	Mismatches:	134
Best Local Similarity:	20.45%	Indels:	49
Query Match:	4	Gaps:	11
DB:			
US-10-698-228-12 (1-1053) x US-09-252-991A-19358 (1-476)			
Ov	4	GGGACCCGGAAAGCCGGAAGCGCGCGGCTGGATCAGGATG	

[illegible]

```

Db      454 LysLysGluLysAsnPheGluThrArgValIleGluTyrGln 468
RESULT 15
US-09-198-452A-1059
; Sequence 1059, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griflais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides,
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 1059
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-1059

Alignment Scores:
Pred. No.:      3.98e-31      Length:      365
Score:          335.00      Matches:     84
Percent Similarity: 46.84%      Conservative: 64
Best Local Similarity: 26.58%      Mismatches:  134
Query Match:    18.15%      Indels:      34
DB:              4          Gaps:         6

US-10-698-228-12 (1-1053) x US-09-198-452A-1059 (1-365)

```

US-10-698-228-12 (1-1053) x US-09-198-452A-1059 (1-365)

QY	130	ATCTTTCCAAATCCAGTACCCCTGATATTTGGAAAATGTATAAACAGGCACAGGCTTCTCTTC	189
Db	50	LeuValProIleLysTyrLysTrpAlaTrpGluHisTyrLeuAsnGlyCysAlaAsn	69
QY	190	TGGACAGACAAGAGGTTGACTTATCAAGAGTATCCCTCACTGG-----AACAGCTT	243
Db	70	TrpLeuProThrGluValProMetAlaArgAspIleGluLeuTrpLysSerAspGluLeu	89
QY	244	AAAGCAGATGAAGACTTCTCATCTCTCACATCTTAGCTTTTTCGACGCCAGTATGGA	303
Db	90	SerGluAspGluArgArgValIleLeuLeuAsnLeuGlyPheSerThrAlaGluSer	109
QY	304	ATTGTAATGAAAATTTGGTGGAGCGCTTAGTCAGGAGGTGCAGGTTCCAGAGGCTCGC	363
Db	110	LeuValGlyAsnAsnIleValLeuAlaIlePheLysHisIleThrAsnProGluAlaArg	129
QY	364	TGTTTCTATGGCTTTCAAATCTCATCGAGAATGTTCACCTCAGAGATGACAGTTGTCTG	423
Db	130	GlnTyrLeuLeuArgGlnAlaPheGluGluAlaValHisThrHisThrPheLeuTyrIle	149
QY	424	ATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAAATGCAATTGAAACC	483
Db	150	CysGluSerLeuGlyLeuAspGluGlyGlu-----ValPheAsnAlaTyrAsnGlu	166
QY	484	ATGCCCTATGTTAAGAAAAGACAGATGTGGCGCTTCGATGGATAGCAGAT-----	534
Db	167	ArgAlaSerIleArgAlaLysAspAspPheGlnMetThrLeuThrValAspValLeuAsp	186
QY	535	-----AGAAAATCTACTTTTGGG	552
Db	187	ProAsnPheSerValGlnSerSerGluGlyLeuGlyGlnPheIleLysAsnLeuValGly	206
QY	553	GAAGAGTGTGGCGCTTGTCTGTGAGAAGAGTTTCTTCTCAGGATCTTTTGTCTGTCT	612
Db	207	TyrTyrIleIle-----MetGluGlyIlePhePheTyrSerGlyPheValMet	222
QY	613	ATATTCCTGGCTAAAGAAGAGAGGTCTTATGCCAGGACTCACTTTTTCCAATCAACTCATC	672
Db	223	IleLeuSerPheHisArgGlnAsnIysMetThrGlyIleGlyGluGlnTyrGlnTyrIle	242
QY	673	AGCAGAGATGAAGGACTTTCATCTGTCACTTGTCTGCTTCTGCTGATGTCTTCCAATAC	732

```
Db      243 LeuArgAspGluThrIleHisLeuAsnPheGlyIleAspLeuIleAsnGlyIleLysGlu 262
QY      733 AAGCCTTCAGAA-----GAAAGGTCAGGAGATCATTGTTGAT 771
Db      263 GluAsnProGluValTrpThrGluLeuGlnGluIleValAlaLeuIleGluLys 282
QY      772 GCTGTCAAAATTGAGCAGGAGTTTTAAACAGAGCCCTGCCAGTTGGCCCTCATTGGAATG 831
Db      283 AlaValGluLeuGluIleGluTyrAlaLysAspCysLeuProArgGlyIleLeuGlyLeu 302
QY      832 AATTGCATTTTGTGAACAGTACATTGAGTTGTAGCTGACAGATTACTTGTGGAACTT 891
Db      303 ArgSerSerMetPheIleAspTyrValArgHisIleAlaAspArgLeuGluArgIle 322
QY      892 GGATTCTCAAAGGTTTTCAGGCAGAAATCCTTTTGATTATG--GAAAACATTCT 948
Db      323 GlyLeuLysProIleTyrHisSerArgAsnProPheProTrpMetSerGluThrMetAsp 342
QY      949 TTAGAAGGAAACAAATTTCTTGAGAAACGAGTTTCAGAGTATTCAG 996
Db      343 LeuAsnLysGluLysAsnPhePheGluThrArgValThrGluTyrGln 358
```

Search completed: October 30, 2005, 06:36:21
Job time : 24.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 30, 2005, 06:34:06 ; Search time 65.75 Seconds
(without alignments)
13390.093 Million cell updates/sec

Title: US-10-698-228-12

Perfect score: 1846

Sequence: 1 atggggaccggaaagcc.....tcaccttgatgcagattttt 1053

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 1865214 seqs, 418043040 residues

Total number of hits satisfying chosen parameters: 3730428

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame*_n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10698228/runat_26102005_100610_4101/app_query.fasta_1.2446
-DB=Published Applications AA -QWTF=fastan -SUFFIX=n2p.rapb -MINMATCH=0.1
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEARSIZE=500 -MINLEN=0
-NCPU=6 -ICPU=3 -NO MAP -LARGESCORE=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA:

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US10F_PUBCOMB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1821	98.6	351	15	US-10-698-228-1
2	1454	78.8	389	15	US-10-698-228-5
3	1454	78.8	389	16	US-10-408-765A-559
4	1454	78.8	389	16	US-10-733-878-457
5	1454	78.8	389	18	US-10-756-149-4899
6	1454	78.8	413	9	US-09-925-301-1347
7	1454	78.8	453	18	US-10-220-335-686
8	1454	78.8	453	18	US-10-450-763-45646
9	1425.5	77.2	386	15	US-10-403-571-76
10	1348	73.0	393	20	US-11-097-143-19182
11	1242	67.3	430	15	US-10-369-493-4122
12	1230	66.6	405	14	US-10-128-714-8245
13	1214	65.8	381	15	US-10-369-493-5686
14	1192.5	64.6	381	14	US-10-128-714-3245
15	1192	64.6	413	14	US-10-032-585-7537
16	1176.5	63.7	321	15	US-10-369-493-10365
17	1169.5	63.4	399	15	US-10-369-493-22170
18	1161.5	62.9	391	15	US-10-369-493-22565
19	1161.5	62.9	401	15	US-10-320-797-3145
20	1147.5	62.2	329	15	US-10-424-599-215866
21	1133.5	61.4	717	16	US-10-425-115-215845
22	1132.5	61.3	345	16	US-10-437-963-113592
23	1128.5	61.1	339	16	US-10-767-701-45287
24	1125.5	61.0	339	16	US-10-425-115-247984
25	1114.5	60.4	329	15	US-10-424-599-269685
26	1097.5	59.5	339	15	US-10-424-599-183313
27	1092.5	59.2	400	16	US-10-425-115-200125
28	1061	57.5	347	16	US-10-739-930-6922
29	916	49.6	270	18	US-10-220-335-342
30	871.5	47.2	344	15	US-10-369-493-13401
31	840	45.5	254	16	US-10-425-115-215843
32	805	43.6	345	15	US-10-369-493-21987
33	784	42.5	225	16	US-10-425-115-215840
34	701	38.0	138	15	US-10-264-237-2195
35	680	36.8	335	18	US-10-450-763-45645
36	662	35.9	181	18	US-10-450-763-37219
37	564.5	30.6	338	14	US-10-156-761-10564
38	544	29.5	156	15	US-10-424-599-215870
39	491	26.6	307	15	US-10-369-493-19363
40	456.5	24.7	145	16	US-10-767-701-45286
41	397	21.5	398	15	US-10-282-122A-51357
42	384.5	20.8	415	15	US-10-282-122A-66252
43	384	20.8	403	15	US-10-282-122A-50118
44	383.5	20.8	415	15	US-10-282-122A-69814
45	380.5	20.6	416	15	US-10-282-122A-68043

ALIGNMENTS

RESULT 1

US-10-698-228-1
; Sequence 1, Application US/10698228
; Publication No. US20040072253A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619W00P
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 1

Qy	841	TTGATGAACACGTACATTGAGTTTGTAGCTGCAGAGATTACTTGTGGAACTTGCAATCTCTCA	900
Db	281	LeuMetLysGlnFyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer	300
Qy	901	AAGCTTTTTCAGGCAGAAAATCCCTTTTGATTTTATGGAAAAACATTTCTTTAGAAAGGAAAA	960
Db	301	LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys	320
Qy	961	ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGGTTATGGCAGAAACC	1020
Db	321	ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr	340
Qy	1021	ACAGATAACGCTCTTCACCTTGGATGCAGATTTT	1053
Db	341	ThrAspAsnValPheThrLeuAspAlaAspPhe	351
RESULT 2			
US-10-698-228-5			
; Sequence 5, Application US/10698228			
; Publication No. US2004007253A1			
; GENERAL INFORMATION:			
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA			
; TITLE OF INVENTION: New Protein and its DNA			
; FILE REFERENCE: 2619WO0P			
; CURRENT APPLICATION NUMBER: US/10/698,228			
; PRIORITY FILING DATE: 2003-10-30			
; PRIOR APPLICATION NUMBER: US/10/019,733			
; PRIOR FILING DATE: 2001-12-28			
; PRIOR APPLICATION NUMBER: JP 11-181131			
; PRIOR FILING DATE: 1999-06-28			
; PRIOR APPLICATION NUMBER: JP 11-192391			
; PRIOR FILING DATE: 1999-07-06			
; PRIOR APPLICATION NUMBER: JP 2000-017770			
; PRIOR FILING DATE: 2000-01-21			
; NUMBER OF SEQ ID NOS: 14			
; SEQ ID NO 5			
; LENGTH: 389			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
; FEATURE:			
; OTHER INFORMATION:			
US-10-698-228-5			
Alignment Scores:			
Pred. No.: 1,96e-147 Length: 389			
Score: 1454.00 Matches: 269			
Percent Similarity: 94.08% Conservative: 33			
Best Local Similarity: 83.80% Mismatches: 19			
Query Match: 78.76% Indels: 0			
DB: 15 Gaps: 0			
US-10-698-228-12 (1-1053) x US-10-698-228-5 (1-389)			
Qy	91	GAAGAGCACCTCTTAAGAAGAGTTCTCGCGGGTTGTTCATCTTCCATCCAGTACCCT	150
Db	69	AspGluProLeuLeuArgGluAsnProArgArgPheValIlePheProIleGluTyrHis	88
Qy	151	GATATTGGAAATGATATAACAGGCACAGGCTTCCTTCGGACAGCAGAGAGTTGAC	210
Db	89	AspIleTrpGlnMetFyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp	108
Qy	211	TTATCAAGAGATCTCCCTCAGTCGAACAGCTTAACAGCATGAGAGTACTTCATCTCT	270
Db	109	LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgFyrPheIleSer	128
Qy	271	CACATCTTAGCCCTTTTTCGAGCAGGTATGGAATTGCTAAATGAAAATTTGGTGGAGCGC	330
Db	129	HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg	148
Qy	331	TTTAGTCAGAGGTGAGGTTCCAGAGGTCGCTGTTTCTATGGCTTCAAATTCATCTCATC	390
Db	149	PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet	168

```
QY 391 GAGAAATGTTCACTACAGATGACAGTTGCTGATAGACACTTACATCAGAGATCCCAAG 450
Db |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
QY 451 AAAGGGAAATTTTATTAATGCAATTAAGAACCATGCCCTATGTTAAGAAAAAGCAGAT 510
Db |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysLysAlaAsp 208
QY 511 TGGGCTTCCTCGATGATAGATGACATGAAATCTACTTTTGGGGAAGAGTGGGCTTT 570
Db |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluValValAlaPhe 228
QY 571 GCTGCTGTAGAGAGTTTCTTCTCAGGATCTTTTGCCTATATCTCGCTATAATCTCGCTAAAGAG 630
Db |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
229 AlaAlaValGluGlyPhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 248
QY 631 AGAGGTCTTATGCCAGGACTCACTTTTCCATGAACTCCTATGATGAGAGATGAAGACTT 690
Db |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluLysLeu 268
QY 691 CACTGTGACTTTGCTGCTGATGTTTCCAACTAGTAAATAGAGCTTCAGAGAAGAG 750
Db |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
QY 751 GTCCAGGAGATCATTTGCTGCTGATGTTTCCAACTAGTAAATAGAGCTTCAGAGAAGAG 810
Db |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
289 ValArgGluIleIleIleAsnAlaValArgIleGluGluGluPheLeuThrGluAlaLeu 308
QY 811 CCAGTTGGCCTCATTTGGAATGCAATTTGATGAAACACAGATTAACCTTTGATGAT 870
Db |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
309 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 328
QY 871 GACAGATTACTTGTGGAATTCGATCTCAAAAGTTTTCAGGAGAAATCTTTTGTAT 930
Db |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348
QY 931 TTTATGGAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACAGAGTTTCAGAG 990
Db |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368
QY 991 TATCAGGCTTTTTCAGTTATGGCAGAAACACACAGATTAACAGTCTTCCACTTGGATGCAGAT 1050
Db |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
369 TyrGlnArgMetGlyValMetSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
QY 1051 TTT 1053
Db |||
389 Phe 389
```

RESULT 3

```
US-10-408-765A-559
; Sequence 359, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 559
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-559
```

Alignment Scores:
Pred. No.: 1.96e-147 Length: 389
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: 16 Gaps: 0

US-10-698-228-12 (1-1053) x US-10-408-765A-559 (1-389)

```
QY 91 GAGAGCCACTCTTAAGAAAGAGTTCTGCGGTTTGTCTATCTTTCCAAATCCAGTACCCT 150
Db |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
69 AspGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGluTyrHis 88
QY 151 GATATTTGGAATGATATAACAGGCACAGCTTCTCTGACACAGCAGAGCTTGAC 210
Db |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
89 AspIleTrpGlnMetTyrLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 108
QY 211 TTATCAAGAGTATCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAACTTACTTCTCT 270
Db |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128
QY 271 CACATCTTAGCCTTTTTCAGCCAGCTGATGGAATTTAAATGAAATTTGGTGGAGCC 330
Db |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
129 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148
QY 331 TTTAGTCAGAGGTGAGGTTCCAGAGCTCGCTGTTCTATGCTTCAAAATTCATCTC 390
Db |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
QY 391 GAGAATCTTCACTCAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAGATCCCAAG 450
Db |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
QY 451 AAAGGGAAATTTTATTAATGCAATTTGCAATTCACCTATGTTTAAAGAAAAAGCAGAT 510
Db |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysLysAlaAsp 208
QY 511 TGGGCTTCCTCGATGATAGATGACATGAAATCTACTTTTGGGGAAGAGTGGGCTTT 570
Db |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluValValAlaPhe 228
QY 571 GCTGCTGTAGAGAGTTTCTTCTCAGGATCTTTTGTCTGCTATATCTGCTGCTAAAGAG 630
Db |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 248
QY 631 AGAGGTCTTATGCCAGGACTCACTTTTCCAACTAGTAAATAGAGCTTCAGAGAAGAG 690
Db |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluLysLeu 268
QY 691 CACTGTGACTTTGCTGCTGATGTTTCCAACTAGTAAATAGAGCTTCAGAGAAGAG 750
Db |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
QY 751 GTCAGGAGATCATTTGTTGATGCTGCTCAAAATGACAGAGGTTTAAACAGAGCCTTG 810
Db |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
289 ValArgGluIleIleIleAsnAlaValArgIleGluGluGluPheLeuThrGluAlaLeu 308
QY 811 CCAGTTGGCCTCATTTGGAATGCAATTTGATGAAACACAGATTAACCTTTGATGAT 870
Db |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
309 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 328
QY 871 GACAGATTACTTGTGGAATTCGATCTCAAAAGTTTTCAGGAGAAATCTTTTGTAT 930
Db |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348
QY 931 TTTATGGAACATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACAGAGTTTCAGAG 990
Db |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368
QY 991 TATCAGGCTTTTTCAGTTATGGCAGAAACACACAGATTAACAGTCTTCCACTTGGATGCAGAT 1050
Db |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
369 TyrGlnArgMetGlyValMetSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
```

```
QY 1051 TTT 1053
Db 389 Phe 389

RESULT 4
US-10-733-878-457
; Sequence 457, Application US/10733878
; Publication No. US20040224408A1
; GENERAL INFORMATION:
; APPLICANT: Jean-Philippe Girard
; APPLICANT: Francois Amalric
; APPLICANT: Myriam Roussigne
; APPLICANT: Thomas Clouaire
; TITLE OF INVENTION: THAP PROTEINS AS NUCLEAR RECEPTORS FOR
; TITLE OF INVENTION: CHEMOKINES AND ROLES IN TRANSCRIPTIONAL REGULATION, CELL
; TITLE OF INVENTION: PROLIFERATION AND CELL DIFFERENTIATION
; FILE REFERENCE: BIOBANK.012A
; CURRENT APPLICATION NUMBER: US/10/733,878
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 60/432699
; PRIOR FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/485027
; PRIOR FILING DATE: 2003-07-03
; NUMBER OF SEQ ID NOS: 535
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 457
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-733-878-457

Alignment Scores:
Pred. No.: 1,96e-147 Length: 389
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: 16 Gaps: 0

US-10-698-228-12 (1-1053) x US-10-733-878-457 (1-389)

QY 91 GAAGGCCACTCTCTAAGAAAGAGTTCTCGCGGTTTGTCTATCTTCCAAATCCAGTACCCT 150
Db 69 AspGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGluTyrHis 88

QY 151 GATATTGGAAATGTPATAACAGGCACAGCGTCTCTCTGGACAGCAGAGGTTGAC 210
Db 89 AspIleTyrGlnMetTyrLysLysAlaGluAlaSerPheTyrThrAlaGluGluValAsp 108

QY 211 TTATCAAGCATCTCCTCACTGGACACAGCTTAAGCAGATGAGAGTACTCTCTCT 270
Db 109 LeuSerLysAspIleGlnHisTyrGluSerLeuLysProGluGluArgTyrPheIleSer 128

QY 271 CACATCTTAGCCCTTTTTCAGCCAGTGTGAATGTAATGAAATTTGGTGGAGCGC 330
Db 129 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148

QY 331 TTTAGTCAGAGGTGCAGGTTCCAGAGGCTCGCTGTCTTCTATGGCTTTCAAAATCTCATC 390
Db 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168

QY 391 GAGATGTTCACTCAGAGATGTACAGTTTGTGTGATGACACTTACATCAGAGATCCCAAG 450
Db 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188

QY 451 AAAAGGGAATTTTATTTAATCAATGAACCATGCCCTATCTTAAAGAAAAGCAGAT 510
Db 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 208

QY 511 TGGGCTTCGCGATGAGATAGACATAAAATCTACTTTTGGGAAAGAGTGGTGGCTTT 570
Db 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValValAlaPhe 228
```

```
QY 571 GCTGCTGTAGACGAGGTTTCTTCTCAGGATCTTTTGTGCTGTATATTCTGGCTAAAGAG 630
Db 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLys 248

QY 631 AGAGGTCTTATGCCAGGACTCAGCTTTTCCAACTCAATCATCAGCAGAGATGAAGACTT 690
Db 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268

QY 691 CACTGTGACTTTGCTTGCCTGATGTTCCAACTACTAGTAATAAGCCTTCAGAAAGG 750
Db 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288

QY 751 GTCAGGAGATCATTTGTGTGCTGTCAAAATTCAGCAGGAGTGTTCCTTAAACAGAACCTTG 810
Db 289 ValArgGluIleIleAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 308

QY 811 CCAGTTGGCTCANTGCAATGCAATTTGATTTGATGAACAGTACATTGAGTTTGTAGCT 870
Db 309 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 328

QY 871 GACAGATTACTTGTGGAACCTTGATTTCTCAAGGTTTTTCAGGCAGAGAAAATCCCTTTGAT 930
Db 329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348

QY 931 TTTATGGAACCAATTTCTTTAGAGGAAACAAATTTCTTTGAGAAACGAGTTTCAGAG 990
Db 349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368

QY 991 TATCAGCGTTTTCAGTTATGCGAGAAACCAACAGATTAACGCTTTCACCTTGGATGCAGAT 1050
Db 369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388

QY 1051 TTT 1053
Db 389 Phe 389

RESULT 5
US-10-756-149-4899
; Sequence 4899, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
; TITLE OF INVENTION: METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/756,149
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4899
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-756-149-4899

Alignment Scores:
Pred. No.: 1,96e-147 Length: 389
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: 18 Gaps: 0

US-10-698-228-12 (1-1053) x US-10-756-149-4899 (1-389)

QY 91 GAAGGCCACTCTCTAAGAAAGAGTTCTCGCGGTTTGTCTATCTTCCAAATCCAGTACCCT 150
Db 69 AspGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGluTyrHis 88

QY 151 GATATTGGAAATGTPATAACAGGCACAGCGTCTCTCTGGACAGCAGAGGTTGAC 210
Db 89 AspIleTyrGlnMetTyrLysLysAlaGluAlaSerPheTyrThrAlaGluGluValAsp 108

QY 211 TTATCAAGCATCTCCTCACTGGACACAGCTTAAGCAGATGAGAGTACTCTCTCT 270
Db 109 LeuSerLysAspIleGlnHisTyrGluSerLeuLysProGluGluArgTyrPheIleSer 128

QY 271 CACATCTTAGCCCTTTTTCAGCCAGTGTGAATGTAATGAAATTTGGTGGAGCGC 330
Db 129 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 148

QY 331 TTTAGTCAGAGGTGCAGGTTCCAGAGGCTCGCTGTCTTCTATGGCTTTCAAAATCTCATC 390
Db 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168

QY 391 GAGATGTTCACTCAGAGATGTACAGTTTGTGTGATGACACTTACATCAGAGATCCCAAG 450
Db 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188

QY 451 AAAAGGGAATTTTATTTAATCAATGAACCATGCCCTATCTTAAAGAAAAGCAGAT 510
Db 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 208

QY 511 TGGGCTTCGCGATGAGATAGACATAAAATCTACTTTTGGGAAAGAGTGGTGGCTTT 570
Db 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValValAlaPhe 228
```

Db 89 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAlaAsp 108
QY 211 TTATCAAGGATCTCCCTCACTGGAAACAAGCTTAAAGCAGATCAGAAAGTACTTTCATCTCT 270
Db 109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 128
QY 271 CACATCTTAGCCCTTTTTCAGCCAGTGTGGAATTTGTAATGAAATTTGGTGGAGCGC 330
Db 129 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAlaValGluArg 148
QY 331 TTATAGTCAGAGGTGCGAGTTCAGAGGCTCGCTGTTCTATGGCTTCAAATTCATC 390
Db 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
QY 391 GAGAAATGTTCACTCAGAGATGTACAGTGTCTGATAGACACTTACATCAGAGATCCCAAG 450
Db 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
QY 451 AAAAGGGAATTTTATTTAATGCAATTGAACCATGCCCTATGTTAAGAAAAAGCAGAT 510
Db 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 208
QY 511 TGGCCCTTCGATGATAGATGAGAAATCTACTTTTGGGAAAGAGTGTGGCTTT 570
Db 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaPhe 228
QY 571 GCTGCTGTAGAAGGATTTCTCTCAGAGTCTTTGCTGCTATATCTGCTTAAAGAAAG 630
Db 229 AlaAlaValGluGlyPhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 248
QY 631 AGAGGTCTTATGTCAGGACTCACTTTTCCAAATGAACTCATCAGCAGAGATGAAGACTT 690
Db 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
QY 691 CACTGTGACTTGTGCTGCTGATGTTCCAAATCTTAAATTAAGCTTTCAGAGAAAGG 750
Db 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
QY 751 GTCAGGAGATCATTTGCTGATGCTGCAAAATGACAGAGTTTTAAACAGAGCTTG 810
Db 289 ValArgGluIleIleAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 308
QY 811 CCAGTTGGCTCATTTGGAATGAATTCATTTTGAATGAAACAGTACATTCAGTTTGTAGCT 870
Db 309 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 328
QY 871 GACAGATTACTTGTGGAATCGATCTCAAAAGGTTTTTCAGCAGAAATCCTTTTGAT 930
Db 329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348
QY 931 TTTATGGAACATTTCTTGAAGGAAACAAATTTCTTTGAGAAACGAGTTTCAGAG 990
Db 349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368
QY 991 TATCAGCGTTTTCAGTATGTCAGAAACCAAGATTAAGTCTTCACTTGGATGAGAT 1050
Db 369 TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
QY 1051 TTT 1053
Db 389 Phe 389

RESULT 6

US-09-925-301-1347
; Sequence 1347, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882

; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1347
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1347

Alignment Scores:
Pred. No.: 2,01e-147 Length: 413
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: Gaps: 0

US-10-698-228-12 (1-1053) x US-09-925-301-1347 (1-413)

QY 91 GAAGGCCACTCTTAAGAAAGAGTTCTCGCGGTTTGTCTATCTTCCAAATCAGTACCT 150
Db 93 AspGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGluTyrHis 112
QY 151 GATATTGGAAATGTATAACAGGCACAGCTTCTCTCGACAGCAGAGAGCTTGCAC 210
Db 113 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 132
QY 211 TTATCAAGGATCTCCCTCACTCGAAACAAGCTTAAAGCAGATGAGAACTTCTATCTCT 270
Db 133 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer 152
QY 271 CACATCTTAGCCCTTTTTCAGCCAGTGTGGAATTTGTAATGAAATTTGGTGGAGCGC 330
Db 153 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 172
QY 331 TTATAGTCAGAGGTGCGAGTTCAGAGGCTCGCTTCTATGGCTTCAAATTCATC 390
Db 173 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 192
QY 391 GAGAATGTTTCACTCAGAGATGTACAGTGTGCTGATAGACACTTACATCAGAGATCCCAAG 450
Db 193 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 212
QY 451 AAAAGGGAATTTTATTTAATGCAATTCGAAACCACTATGCCCTATGTTAAGAAAAAGCAGAT 510
Db 213 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysAlaAsp 232
QY 511 TGGCCCTTCGATGATAGCAGATAGAAATCTACTTTTGGGAAAGAGTGTGGCTTT 570
Db 233 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaPhe 252
QY 571 GCTGCTGTAGAAGGATTTTCTCTCAGGATCTTTGCTGCTATATCTGCTGCTAAAGAG 630
Db 253 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 272
QY 631 AGAGGTCTTATGTCAGGACTCACTTTTCCAAATGAACTTCATCAGCAGAGATGAAGACTT 690
Db 273 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 292
QY 691 CACTGTGACTTGTGCTGCTGATGTTCCAAATGAACTTAAATGAACTTCCAGAGAAAGG 750
Db 293 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 312
QY 751 GTCAGGAGATCATTTGATGCTGCAAAATTCAGCAGAGAGTTTTAAACAGAGAGCTTG 810
Db 313 ValArgGluIleIleIleAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 332
QY 811 CCAGTTGGCTCATTTGGAATGAATTTGATTTGATGAAACAGATGACATTCAGTTTGTAGCT 870
Db 333 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 352

Qy	871	GACAGATTACTTGTGGAAC	TCTGGATTCTCAAGGTTTTTTCAGGCAGAAAATCCTTTGAT	930
Db	393	AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp	412	
Qy	931	TTTATGGAAACATTTCTTTAGAGGAAAACAAATTTCTTTGAGAAACAGGTTTCAGAG	990	
Db	413	PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu	432	
Qy	991	TATCAGGTTTTGCAGTTATGGCAGAAAACACAGATAACGTCCTTCACCTTGGATCGAGAT	1050	
Db	433	TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp	452	

Db 433 TYGINargmetGlyvalmetSerProInrGluAsnSerPheTrnLeuASPAlaAsp 452
 Qy 1051 TTT 1053
 Db 453 Phe 453

; sequence %6, Application US/10403571
 ; Publication No. US20040068763A1

```

; GENERAL INFORMATION:
; APPLICANT: Hopkins, Nancy
; APPLICANT: Golling, Gregory
; APPLICANT: Amsterdam, Adam
; APPLICANT: Sun, Zhaoxia
; TITLE OF INVENTION: Developmental Mutations in Zebrafish
; FILE REFERENCE: 01997/539002
; CURRENT APPLICATION NUMBER: US/10/403,571
; CURRENT FILING DATE: 2003-03-25
; PRIOR APPLICATION NUMBER: US 60/368,760
; PRIOR FILING DATE: 2002-03-29
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 76
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Danio rerio
US-10-403-571-76

Alignment Scores:
Pred. No.:      2,34e-144      Length:      386
Score:          1425.50      Matches:    268
Percent Similarity: 89.41%    Conservative: 36
Best Local Similarity: 78.82%  Mismatches: 31
Query Match:      77.22%     Indels:     5
DB:               15         Gaps:       1

US-10-698-228-12 (1-1053) x US-10-403-571-76 (1-386)

QY      49  AGATCATCTTCAGACACCAACGAAGTGAATAAGTCA-----AATGAA  93
      |||      |||      |||      |||      |||      |||
Db      47  ArgLysIlePheAspGluSerGluGlyGlnSerLysAlaLysLysGlyAlaValGluGlu  66

QY      94  GAGCCACTCTTAAGAAAGAGTTCTCGCCGGTTGTTCATCTTTCCAAATCCAGTACCCGTGAT  153
      |||      |||      |||      |||      |||      |||
Db      67  GluProLeuLeuLysGluAsnProHisArgPheValIlePheProIleGlnTyrHisAsp  86

QY      154  ATTTGGAAATGTATAACAGCACAGCAGGCTTCCTTCTGGACAGCAGAGAGGTGACTTGA  213
      |||      |||      |||      |||      |||      |||
Db      87  IleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpTrpAlaGluGluValAspLeu  106

QY      214  TCAAAGGATCTCCCTCAGTCGGAACAGCTTAAGCAGATGAGAAGTACTTCTCTCTCAC  273
      |||      |||      |||      |||      |||      |||
Db      107  SerLysAspLeuGlnHisTrpAspSerLeuLysAspGluGluArgTyrPheIleSerHis  126

QY      274  ATCTTAGCCCTTTTTCAGCCAGTGAATGAAATTGTAATGAAAAATTGGTGAGGCGCTTT  333
      |||      |||      |||      |||      |||      |||
Db      127  ValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArgPhe  146

QY      334  AGTCAGGAGGTGAGGTTCCAGAGGCTCGCTGTTCTATATGGCTTTTCAAAATCTCATCGAG  393
      |||      |||      |||      |||      |||      |||
Db      147  ThrGlnGluValGlnValThrGluAlaAcGysPheTyrGlyPheGlnIleAlaMetGlu  166

```

```
QY 394 AATGTTCTACTAGAGATGTACAGTTTGTCTGATAGACACTTACATCAGAGATCCCAAGAAA 453
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
167 AsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspSerLysGlu 186
QY 454 AGGAATTTTATTAATCAATTGAACCATGCGCTATGTTAAGAAAAAAGCAGATTGG 513
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
187 ArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysLysAlaAspTrp 206
QY 514 GCCTTCGATGATGACAGATGAAATCTACTTTTGGGGAAGAGTGGCTTGTCT 573
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
207 AlaLeuAsnTrpIleGlyAspLysAsnAlaArgTyrGlyGluArgValValAlaPheAla 226
QY 574 GCTGTAGAAGAGTCTTCTCTCAGAGATCTTTGCTGCTATATCTCGGCTTAAGAGAGA 633
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
227 AlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLysArg 246
QY 634 GGTCTTATGCCAGGACTCACTTTTCCAAATGAACCTCATCAGCAGAGATGAAGGACTTCAC 693
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
247 GlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHis 266
QY 694 TGTGATTTTGTCTGCTGATGTTCCAAATCTAGTAAATGAAGCTTCAGAGAAAGGCTC 753
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
267 CysAspPheAlaCysLeuMetPheLysHisLeuIleAsnLysProSerGluGluThrVal 286
QY 754 AGGAGATCATTTGTTGATGCTGCTCAAAATTTGAGCAGAGTTTTAAACAGAGCTTGCCA 813
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
287 LysLysIleIleMetAsnAlaValGluIleGluGluPheLeuThrAspAlaLeuPro 306
QY 814 GTTGGCTCATTTGAATGAATTCATTTTGAATGAACACAGTACATTCAGTTTGTAGCTGAC 873
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
307 ValLysLeuIleGlyMetAsnCysAspLeuMetLysGlnTyrIleGluPheValAlaAsp 326
QY 874 AGATTACTTGTGGAATTCGATCTCAAGGTTTTCAGGCAGAAATCTTTTGTATTTT 933
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
327 ArgLeuLeuLeuGluLeuGlyPheAspLysValTyrArgValGluAsnProPheAspPhe 346
QY 934 ATGGAAACATTTCTTTAGAAGAAAAACAAATTTCTTTGAGAAACGAGTTTCAGAGTAT 993
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
347 MetGluAsnIleSerLeuGluGlyLysThrAsnPheGluLysArgValGlyGluTyr 366
QY 994 CAGCGTTTTCAGTTATGCGCAAAACACAGATAACGTCTTCACTTGGATGCGATTTT 1053
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
367 GlnArgMetGlyValMetSerGlyThrThrAspAsnThrPheThrLeuAspAlaAspPhe 386

RESULT 10
US-11-097-143-19182
; Sequence 19182, Application US/11097143
; Publication No. US2005020858A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; PRIOR FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
```

```
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19182
; LENGTH: 393
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-19182

Alignment Scores:
Pred. No.: 5,47e-136 Length: 393
Score: 1348.00 Matches: 252
Percent Similarity: 85.92% Conservative: 41
Best Local Similarity: 73.90% Mismatches: 48
Query Match: 73.02% Indels: 0
DB: 20 Gaps: 0

US-10-698-228-12 (1-1053) x US-11-097-143-19182 (1-393)
QY 31 GGGCTGGATCAGGATGAGAGATCATCTTCAGACACCAACGAAAGTAAATAAGTCAAT 90
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
53 GlyIleGlyLysSerAlaAsnSerLeuMetGluLysSerValThrProPheAspProSer 72
QY 91 GAAGAGCCACTCTTAAGAAAGAGTTCTGCGCGTTTGTCTATCTTTCCAAATCCAGTACCCT 150
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
73 LeuGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGlnTyrHis 92
QY 151 GATATTTGGAAATGTATAACAGGCACAGGCTCTCTCGACAGCAGAGAGGTTGAC 210
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
93 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrValGluGluValAsp 112
QY 211 TTATCAAAAGGATCTCCCTCTAGCAACAGCTTAAAGCAGATGAGAGTACTTCTATCTCT 270
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
113 LeuSerLysAspLeuThrAspTrpHisArgLeuLysAspAspGluArgHisPheIleSer 132
QY 271 CACATCTTAGCCCTTTTTCAGCAGCATGATGGAATGTAAATGAAATTTGGTGAGCGC 330
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
133 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg 152
QY 331 TTTAGTCAGCAGGTCGAGGTTCCAGAGGCTCGCTGTTCTATGCTTTTCAATTTCTCATC 390
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
153 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 172
QY 391 GAGAATGTTCACTCAGAGATGTACAGTTGCTGTAGTACACTTACATCAGAGATCCCAAG 450
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
173 GluAsnValHisSerGluMetTyrSerValLeuIleAspThrTyrIleArgAspProHis 192
QY 451 AAAAGGGAATTTTATTTAATGCAATTCATGCAATTCATCTTTGGGGAAGAGTGGTGCCTTT 570
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
193 GlnArgGluTyrLeuPheAsnAlaIleGluThrMetProAlaValLysArgLysAlaAsp 212
QY 511 TGGGCTTGGATGGATGATAGATAGAAATCTACTTTTGGGGAAGAGTGGTGCCTTT 570
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
213 TrpAlaLeuSerTrpIleSerSerLysSerAlaAsnPheGlyGluArgIleAlaPhe 232
QY 571 GCTGCTGTAGAAGAGTTTCTCTCAGGATCTTTTCTGCTGTATATTCTGGCTAAAGAG 630
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
233 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 252
QY 631 AGAGGTCTTATGCGCAGGACTCAGCTTTTCCAAATGAACTCATCAGCAGAGATGAAGACTT 690
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
253 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 272
QY 691 CACTGTGACTTTGCTTGCCTGATGTTCCAAATCTATCTAGTAATAGACCTTCAGAGAAAGG 750
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
273 HisCysAspPheAlaValLeuMetPheGlnHisLeuValGlnArgProLysArgGluArg 292
QY 751 GTCAGGAGAGATCATTTGTTGATGCTGCTCAAAATTCAGCAGGAGTGTAAACAGAGCCTTG 810
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
293 IleIleGluIleIleArgAspAlaValAlaIleGluGlnGluPheLeuThrAspAlaLeu 312
QY 811 CCAGTTGCGCTCATTTGGAATGAATTCATTTTGAATGAACAGTACATTGAGTTGTAGCT 870
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
313 ProValAsnLeuIleGlyMetAsnCysAspLeuMetSerGlnTyrIleGluPheValAla 332
```


; CURRENT FILING DATE: 2002-04-23
 ; PRIOR APPLICATION NUMBER: US 60/285,697
 ; PRIOR FILING DATE: 2001-04-23
 ; PRIOR APPLICATION NUMBER: US 60/287,066
 ; PRIOR FILING DATE: 2001-04-27
 ; PRIOR APPLICATION NUMBER: US 60/295,890
 ; PRIOR FILING DATE: 2001-06-05
 ; PRIOR APPLICATION NUMBER: US 60/303,899
 ; PRIOR FILING DATE: 2001-07-09
 ; PRIOR APPLICATION NUMBER: US 60/316,362
 ; PRIOR FILING DATE: 2001-08-31
 ; NUMBER OF SEQ ID NOS: 8603
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 8245
 ; LENGTH: 405
 ; TYPE: PRT
 ; ORGANISM: Aspergillus fumigatus
 US-10-128-714-8245

Alignment Scores:
 Pred. No.: 3,03e-123 Length: 405
 Score: 1230.00 Matches: 234
 Percent Similarity: 81.95% Conservative: 52
 Best Local Similarity: 67.05% Mismatches: 49
 Query Match: 66.63% Indels: 14
 DB: 14 Gaps: 3

US-10-698-228-12 (1-1053) x US-10-128-714-8245 (1-405)

QY 46 GAGAGATCATCTTCAGACACCAAGAGTGAATAAGTCAATCAAGAGCCACTCTTA 105
 DB 58 GluAlaSerLysValAlaGlnSerIleLysGluLeuGluAlaAsn--GluProLeuLeu 76
 QY 106 AGAAGAGTTCGCGCGTGTCTCATCTTCCAAATCCAGTACCTCGATATTTGGAAAATG 165
 DB 77 GlnGluAsnProHisArgPheValLeuPheProIleLysTyrHisGluIleTrpGlnMet 96
 QY 166 TATAACAGGACAGGCTTCCTCTCGGACAGACAGAGGTGACTTATCAAGAGATCTC 225
 DB 97 TyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluIleAspLeuSerLysAspLeu 116
 QY 226 CTTCACTGG---AACAAGCTTAAGACAGATGAGAAGTACTTCTCATCTCTCAGATCTTAGCC 282
 DB 117 HisAspTrpAsnAsnArgLeuAsnAspAspGluArgTyrPheIleSerHisValLeuAla 136
 QY 283 TTTTTCAGCCAGCATGGAATTTGAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAG 342
 DB 137 PhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuLeuGluArgPheSerAsnGlu 156
 QY 343 GTGAGGTTCCAGAGCTGCTGTTTCTATGCTTCAAAATCTCATCTCAGAGATGTTTAC 402
 DB 157 ValGlnValProGluAlaArgCysPheTyrGlyPheGlnIleMetIleGluAsnIleHis 176
 QY 403 TCAGAGATGTACAGTTTGTGTAGACACTTACATCAGAGATCCCAAGAAAAGGCAATTT 462
 DB 177 SerGluThrTyrSerLeuLeuIleAspThrTyrIleLysGluProLysGlnArgThrTyr 196
 QY 463 TTATTATGCAATTTGAACCAATGCGCTATGTTTAAGAAAAGAGATTTGGCGCTTGCA 522
 DB 197 LeuPheAspAlaIleAspThrValProCysIleArgLysLysAlaGlnTrpAlaMetArg 216
 QY 523 TGGATAGCAGATAGAAAATCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAGAA 582
 DB 217 TrpIleSerAspLysGluSerThrPheAlaGlnArgLeuValAlaPheAlaAlaValGlu 236
 QY 583 GGAGTTTCTCTCAGGATCTTTGCTGTATTTCTGCTAATGCTGCTAAGAGAGAGTCTTATG 642
 DB 237 GlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLysArgGlyLeuMet 256
 QY 643 CCAGGACTCATTCTTCCAAATGACTCATCAGCAGAGATGAAGGACTTCCACTGTGACTTT 702
 DB 257 ProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisThrAspPhe 276

QY 703 GCTTCCTGATGTTCCTCAATACCTTAGTAATAAGCCCTTCAGAAGAAAGGTCAGGAGATC 762
 DB 277 AlaCysLeuLeuPheSerHisLeuAsnHisArgProAspLysLysValIleGluAspIle 296
 QY 763 ATTGTGATGCTGTCAAAATTTGAGCAGGAGTCTTTTAAACAGAGCTTGCAGTTGGCCTC 822
 DB 297 IleValGluAlaValAlaIleGluGlnGluPheLeuThrAspAlaLeuProValAlaLeu 316
 QY 823 ATTGAATGAATTCGATTTTGTATGAACACGATACATTCAGTTTGTAGCTGCAGATTACTT 882
 DB 317 LeuGlyMetAsnSerLysLeuMetCysGlnTyrIleGluPheValAlaAspArgLeuLeu 336
 QY 883 GTGGAATTCGATTCCTCAAGGTTTTCAGGCAAGAAATCCTTTTGTATTTATGGAAC 942
 DB 337 ValAlaLeuGlyAsnLysTyrPheAsnSerProAsnProPheAspPheMetGluSer 356
 QY 943 ATTCTTTAGAGAGAAAACAAATTTCTTTGAGAAAACGAGTTTCAGAGTATCAGCGTTT 1002
 DB 357 IleSerLeuAlaGlyLysThrAsnPheGluLysArgValGlyAspTyrGlnLysAla 376
 QY 1003 GCAGTTTATGCGCAGAAACC-----ACAGAT 1026
 DB 377 GlyValMetAlaSerThrLysLysAspAlaSerGlnAspThrAlaLysAspAlaAsnAsp 396
 QY 1027 AACGTCTTCACCTTGGATGCGAGATTTT 1053
 DB 397 GlyGlyLeuCysPheAspGluAspPhe 405

RESULT 13
 US-10-369-493-5686
 ; Sequence 5686, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 5686
 ; LENGTH: 381
 ; TYPE: PRT
 ; ORGANISM: Caenorhabditis elegans
 US-10-369-493-5686

Alignment Scores:
 Pred. No.: 1.58e-121 Length: 381
 Score: 1214.00 Matches: 231
 Percent Similarity: 81.10% Conservative: 48
 Best Local Similarity: 67.15% Mismatches: 61
 Query Match: 65.76% Indels: 4
 DB: 15 Gaps: 2

US-10-698-228-12 (1-1053) x US-10-369-493-5686 (1-381)

QY 34 CTGGATCAGATCAGATCATCTTCAGAC-----ACCAACGAAGTGAATAAG--- 84
 DB 38 ValAspGlnThrLysAlaAlaSerAlaGluGluThrAsnAsnGluSerGluValAsnGlu 57
 QY 85 ---TCAATGAAGAGCCACTCTTAAGAAAGAGTCTCGCGGTTTGTTCATCTTCCCAATC 141
 DB 58 LeuAspAlaAspGluProMetLeuGlnAspLeuAspAsnArgPheValIlePheProLeu 77
 QY 142 CAGTACCTCATATTTGGAAAATGTATAACACGACAGCGTCTCTCTGACAGCAGAA 201
 DB 78 LysHisHisAspIleTrpAsnPheTyrLysLysAlaValAlaSerPheTrpThrValGlu 97

QY	202	GAGGTTGACTTATCAAAAGGATCTCCCTCAC	TGGACAGCGCTTAAAGCAGATGAGAGTAC	261
DB	98	GIUValAspLeuGlyAspMetAsnAspTrpGIU	lysMetAsnGlyAspGIUlnTr	117
QY	262	TTTCATCTCTCAATCTTAGCCCTTTTTTG	CAGCCAGTGATGAATTTGTAATAAGAAATTTG	321
DB	118	PhelIserArgIleLeuAlaPhePheAla	AspGlyIleValAsnGIUAsnLeu	137
QY	322	GTGAGAGCGCTTTAGTCAGGAGGTGCAGG	TTCACAGGCTCGCTGTTCATATGGCTTTTCAA	381
DB	138	CysGIUArgPheSerAsnGIUValGlnVal	SerGIUAlaAArgPhePheTyrglyPheGln	157
QY	382	ATTCTCATCGAGATGTTCTACTCAGAGAT	GTACAGTTTGTGTCATAGACACTTACATCAGA	441
DB	158	IleAlaIleGIUAsnIleHisSerGIUmet	TyrSerLysLeuIleGIUlnTrTyrIleArg	177
QY	442	GATCCCAAGAAAGGGGAATTTTATTAA	TTCGAATTTGAAACCATCGCCCTATGTTAAAGAAA	501
DB	178	AspGIUlnTrGIUArgAsnThrLeuPheAsn	AlaValAspGIUPheIleGlyLys	197
QY	502	AAAGCAGATTGGCGCTTCGAGTACGAT	ACAGATGAAATACTATTTTGGGGAAGAGTG	561
DB	198	LysAlaAspTrpAlaLeuArgTrpIleSer	AspLysLysAlaSerPheAlaGIUArgLeu	217
QY	562	GTGGCCTTTCGTGCTGTAGAGGAGTTTCT	TCTCAGGATCTTTTGTGCTATATTTCTGG	621
DB	218	IleAlaPheAlaAlaValGIUArgIlePhe	PheSerGlySerPheAlaSerIlePheTrp	237
QY	622	CTAAGAAGACAGAGCTTTATGCCAGGAC	TCACTTTTCCAAATGAATCACTCAGCAGAGAT	681
DB	238	LeuLysLysArgGlyLeuMetProGIU	leuThrHisSerAsnGIUleuIleSerArgAsp	257
QY	682	GAAGGACTTCACGTGTGCTTTGCTGTG	CTGTGCTGTATTCCAATCTTAGTAATAAGCCCTTCA	741
DB	258	GIUGlyLeuHisArgAspPheAlaCys	LeuLeuTyrSerLysLeuGlnLysLeuThr	277
QY	742	GAAGAAAGGTCAGGAGATCATTTGTATG	CTCTGTCAAAATTGACGAGGAGTTTTTAAACA	801
DB	278	GlnGlnArgIleTyrAspIleIleLys	AspAlaValAlaIleGIUlnGlnPheLeuThr	297
QY	802	GAAGCCTTGCCAGTTGGCCTCATTTGGA	ATGAATTTGATGAAACAGTACATTGAG	861
DB	298	GIUAlaLeuProValAspMetIleGly	MetAsnCysArgLeuMetSerGlnTyrIleGlu	317
QY	862	TTTCTAGCTGACAGATTACTTGTGGAA	CTTGTGGATTTCTCAAAGGTTTTTCAGGCGAATAA	921
DB	318	PheValAlaAspHisLeuLeuValGIU	leuGlyCysAspLysLeuTyrLysSerLysAsn	337
QY	922	CCTTTTGATTTATGGAAACATTTCTT	TAGAAGGAAACAAATTTCTTTGAGAAACGA	981
DB	338	ProPheAspPheMetGIUAsnIleSer	IleAspGlyLysThrAsnPhePheGIULysArg	357
QY	982	GTTTTCAGGTATCAGCGCTTTTTCAG	CTTTATGGCAGAAACCCACAGATAAGCTCTTCCACCTTG	1041
DB	358	ValSerGIUlnTrGlnArgProGIU	ValMetValAsnGIUAlaGIUArgGlnPheAspLeu	377
QY	1042	GATCCAGATTTT	1053	
DB	378	GIUAlaAspPhe	381	

RESULT 14

```

RES001 14
US-10-128-714-3245
; Sequence 3245, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wenqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and

```

```
Db 233 ProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeuHisThrAspPhe 252
QY 703 GCTTGCCTGATGTTCCCAATAGTAAATAGCCTTCAGAGAAAGGTCAGGAGATC 762
Db 253 AlaCysLeuLeuPheSerHisLeuAsnHisArgProAspLysLysValIleGluAspIle 272
QY 763 ATTGTTGATGCTGTCACAAATTCAGCAGGAGTTCCTTAAACAGAACGCTTCGAGTTGCCCTC 822
Db 273 IleValGluAlaValAlaIleGluGlnGluPheLeuThrAspAlaLeuProValAlaLeu 292
QY 823 ATTGAATGAATGCAATTTTGATGAACAGTACATGATGTTGTAGCTGACAGATTACTT 882
Db 293 LeuGlyMetAsnSerLysLeuMetCysGlnTyrIleGluPheValAlaAspArgLeuLeu 312
QY 883 GTGGAACCTGGATTCCTCAAGTTCCTTTCAGGACAGAAATCCTTTGATTTATGGAAC 942
Db 313 ValAlaLeuGlyAsnLysLysTyrPheAsnSerProAsnProPheAspPheMetGluSer 332
QY 943 ATTTCTTTAGAGAAACAAATTTCTTTGAGAACGAGTTTCAGAGTATCAGCGCTTTT 1002
Db 333 IleSerLeuAlaGlyLysThrAsnPhePheGluLysArgValGlyAspTyrGlnLysAla 352
QY 1003 GCAGTTATGGCAGAAACC-----ACAGAT 1026
Db 353 GlyValMetAlaSerThrLysLysAspAlaSerGlnAspThrAlaLysAspAlaAsnAsp 372
QY 1027 AACGTCTTCACCTTGATGACAGATTTT 1053
Db 373 GlyGlyLeuCysPheAspGluAspPhe 381
RESULT 15
US-10-032-585-7537
; Sequence 7537, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; FILE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7537
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-032-585-7537
Alignment Scores:
Pred. No.: 3,86e-119 Length: 413
Score: 1192.00 Matches: 225
Percent Similarity: 80.17% Conservative: 58
Best Local Similarity: 63.74% Mismatches: 64
Query Match: 64.57% Indels: 6
DB: 14 Gaps: 3
US-10-698-228-12 (1-1053) x US-10-032-585-7537 (1-413)
QY 13 GAAAGCGCGAAGCGCGCGGTGGATCAGGATCAGATCACTTCAGACACCAACGAA 72
Db 61 GluLysAspSerLeuLysLysHisGlnAspPheLeuAlaLysHisLysValHisArg 80
QY 73 AGTGAATAAAG-----TCAATGAGAGCCACTCCTTAAGAAAGAGTTCTCGCCGTTT 126
Db 81 HisLysLeuLysGlnLeuGluAlaGluProLeuValGluAsnLysArgTyr 100
QY 127 GTCATCTTTCCCAATCCAGTACCTGATATTTGGAAAATGTATAAACAGCAGCAGGCTTC 186
Db 101 ValMetPheProIleArgTyrHisGluIleTyrAsnPheTyrLysLysAlaGluAlaSer 120
```

```
QY 187 TTCTGCACACAGAGAGGTTGACTTATCAAGGATCTCCCTCACTGG- - -AACAGCTT 243
Db 121 PheTyrThrAlaGluGluIleAspLeuSerLysAspLeuAspTyrPheAsnLysLeu 140
QY 244 AAACAGATCAGAGTACTTCTCTCATCTCTCAGATCTTAGCCTTTTTCAGCAGGATCGA 303
Db 141 AsnGluAsnGluArgTyrPheIleSerArgValLeuAlaPhePheAlaAspSerGly 160
QY 304 ATTGTAATAAATTTGGTGGAGCGCTTTAGTCAGAGGTCAGAGGTCAGAGGCTCGC 363
Db 161 IleValGlyGluAsnLeuIleGluAsnPheSerThrGluValGlnLeuProGluAlaLys 180
QY 364 TGTTCCTATCGCTTTCAAAATCTCATCGAGAATCTTCACTCAGAGATCTACAGTTTGTG 423
Db 181 SerPheTyrGlyPheGlnIleMetMetGluAsnIleHisSerGluThrTyrSerLeuLeu 200
QY 424 ATAGACACTTATCATCAGAGATCCCAAGAAAGGAATTTTATTATTAATGCAATGAAACC 483
Db 201 IleGluThrTyrIleLysAspProGlnGluAlaAspTyrLeuPheAsnAlaIleAlaAsn 220
QY 484 ATGCCCTATGTTAAGAAAAAGCAGATGGCCCTTGCATCGATCGATAGCAGATAGAAAATCT 543
Db 221 IleProCysIleGlnLysLysAlaAspTyrAlaIleLysTyrIleGlnAspAspGluAla 240
QY 544 ACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTCTAGAGAGGATTTTCTTCTCAGGATCT 603
Db 241 LeuTyrGlyGluArgLeuValAlaPheAlaAlaValGluGlyIlePhePheSerGlySer 260
QY 604 TTTGCTGCTATATTCTGGCTTAAAGAGAGAGGCTTATGCCAGGACTCACATTTTCCAAT 663
Db 261 PheAlaSerIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSerAsn 280
QY 664 GAATCATCAGCAGATGAAGGACTTCACCTGCTGCTAGTCTGCTGCTGCTGCTGCTGCTGCT 723
Db 281 GluLeuIleCysArgAspGluGlyLeuHisThrAspPheAlaCysLeuLeuPheAlaHis 300
QY 724 TTAGTAAATAAGCTTCAGAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAAAATT 783
Db 301 LeuGlnAsnArgProSerProGluIleValGluArgIleIleThrGluAlaValAspVal 320
QY 784 GAGCAGAGTGTTTTAAACAGAGCGCTTCGCGCTTCGCTCATTTGGAATGAATTTGCTTTTG 843
Db 321 GluLysGluTyrPheThrAspAlaLeuProValSerLeuLeuGlyMetAsnCysAsnLeu 340
QY 844 ATGAAAACAGTACATTTGAGTTGTAGCTGACAGATTAATTTGTTGAGAACTTCTCAAG 903
Db 341 MetCysGlnTyrIleGluPheValAlaAspArgLeuLeuLeuAlaLeuGlyAsnLysLys 360
QY 904 GTTTTTCAGCAGCAAAATCCTTTTATTTATGGAACCAATTTCTTTAGAGGAAACACA 963
Db 361 ValTyrAsnValThrAsnProPheAspPheMetGluAsnIleSerLeuAlaGlyLysThr 380
QY 964 AATTTCTTTCAGAAACGAGTTTCAGAGTATCAGCGTTCGAGTTCGAGTATGCGCAGAACCA 1023
Db 381 AsnPhePheGluLysArgValSerAspTyrGlnLysAlaGlyValMetAlaLysThrGlu 400
QY 1024 GAT-----AACGTCTTTCACCTTCGATCGACAGATTTT 1053
Db 401 AsnLysGluAlaAspAlaPheThrPheAspGluAspPhe 413
```

Search completed: October 30, 2005, 06:55:20
Job time : 80.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 30, 2005, 06:24:21 ; Search time 173.25 Seconds
(without alignments)
14198.102 Million cell updates/sec

Title: US-10-698-228-12
Perfect score: 1846
Sequence: 1 atggcgaccggaagcc.....tcaccttgatgcagatttt 1053

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 6959266 seqs, 1168006243 residues

Total number of hits satisfying chosen parameters: 13918532

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlh
-MODE=frame_n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10698228/runat_26102005_100609_4045/app_query.fasta_1.2446
-DB=Pending Patents AA Main -QFMT=fastan -SUFFIX=n2p.rapm -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEARSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10698228 -cgn2_1_523 @runat_26102005_100609_4045
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=120 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending Patents AA Main:*

- 1: /cgn2_6/ptodata/1/paa/PCTUS COMB.pap:*
- 2: /cgn2_6/ptodata/1/paa/US06 COMB.pap:*
- 3: /cgn2_6/ptodata/1/paa/US07 COMB.pap:*
- 4: /cgn2_6/ptodata/1/paa/US08 COMB.pap:*
- 5: /cgn2_6/ptodata/1/paa/US081 COMB.pap:*
- 6: /cgn2_6/ptodata/1/paa/US082 COMB.pap:*
- 7: /cgn2_6/ptodata/1/paa/US083 COMB.pap:*
- 8: /cgn2_6/ptodata/1/paa/US084 COMB.pap:*
- 9: /cgn2_6/ptodata/1/paa/US085 COMB.pap:*
- 10: /cgn2_6/ptodata/1/paa/US086 COMB.pap:*
- 11: /cgn2_6/ptodata/1/paa/US087 COMB.pap:*
- 12: /cgn2_6/ptodata/1/paa/US088 COMB.pap:*
- 13: /cgn2_6/ptodata/1/paa/US089 COMB.pap:*
- 14: /cgn2_6/ptodata/1/paa/US090 COMB.pap:*
- 15: /cgn2_6/ptodata/1/paa/US091 COMB.pap:*
- 16: /cgn2_6/ptodata/1/paa/US092 COMB.pap:*
- 17: /cgn2_6/ptodata/1/paa/US093 COMB.pap:*
- 18: /cgn2_6/ptodata/1/paa/US094 COMB.pap:*
- 19: /cgn2_6/ptodata/1/paa/US095 COMB.pap:*
- 20: /cgn2_6/ptodata/1/paa/US096 COMB.pap:*
- 21: /cgn2_6/ptodata/1/paa/US097A COMB.pap:*
- 22: /cgn2_6/ptodata/1/paa/US097B COMB.pap:*
- 23: /cgn2_6/ptodata/1/paa/US098 COMB.pap:*
- 24: /cgn2_6/ptodata/1/paa/US099A COMB.pap:*
- 25: /cgn2_6/ptodata/1/paa/US099B COMB.pap:*
- 26: /cgn2_6/ptodata/1/paa/US100 COMB.pap:*
- 27: /cgn2_6/ptodata/1/paa/US101 COMB.pap:*
- 28: /cgn2_6/ptodata/1/paa/US102 COMB.pap:*

- 29: /cgn2_6/ptodata/1/paa/US103 COMB.pap:*
- 30: /cgn2_6/ptodata/1/paa/US104 COMB.pap:*
- 31: /cgn2_6/ptodata/1/paa/US105 COMB.pap:*
- 32: /cgn2_6/ptodata/1/paa/US106 COMB.pap:*
- 33: /cgn2_6/ptodata/1/paa/US107 COMB.pap:*
- 34: /cgn2_6/ptodata/1/paa/US108 COMB.pap:*
- 35: /cgn2_6/ptodata/1/paa/US109 COMB.pap:*
- 36: /cgn2_6/ptodata/1/paa/US110 COMB.pap:*
- 37: /cgn2_6/ptodata/1/paa/US60 COMB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1821	98.6	351	22	US-09-787-491B-11
2	1821	98.6	351	22	US-09-791-537-77474
3	1821	98.6	351	27	US-10-170-205B-19045
4	1821	98.6	351	32	US-10-698-228-1
5	1821	98.6	351	35	US-10-990-328-7547
6	1821	98.6	351	37	US-60-128-660-7
7	1821	98.6	351	37	US-60-449-629-94
8	1821	98.6	351	37	US-60-505-218-322
9	1821	98.6	366	22	US-09-791-537-27671
10	1817	98.4	351	20	US-09-629-469A-12811
11	1817	98.4	351	22	US-09-791-537-112185
12	1817	98.4	351	35	US-10-917-503-12811
13	1454	78.8	389	1	PCT-US02-10824-143
14	1454	78.8	389	1	PCT-US03-17409-171
15	1454	78.8	389	10	US-08-663-617-33
16	1454	78.8	389	12	US-08-837-599-33
17	1454	78.8	389	22	US-09-791-537-88179
18	1454	78.8	389	27	US-10-170-205B-8018
19	1454	78.8	389	30	US-10-408-765-559
20	1454	78.8	389	30	US-10-408-765A-559
21	1454	78.8	389	32	US-10-698-228-5
22	1454	78.8	389	33	US-10-733-878-457
23	1454	78.8	389	33	US-10-756-149-4899
24	1454	78.8	389	35	US-10-940-774-6016
25	1454	78.8	389	37	US-60-389-987-559
26	1454	78.8	389	37	US-60-412-418-559
27	1454	78.8	389	37	US-60-452-680-20616
28	1454	78.8	389	37	US-60-568-073-984
29	1454	78.8	413	1	PCT-US00-05882-1347
30	1454	78.8	413	24	US-09-925-301-1347
31	1454	78.8	449	35	US-10-990-328-12349
32	1454	78.8	449	35	US-10-990-328-12350
33	1454	78.8	453	1	PCT-US01-04926A-686
34	1454	78.8	453	1	PCT-US01-08631-45646
35	1454	78.8	453	28	US-10-220-335-686
36	1454	78.8	453	28	US-10-290-752-686
37	1454	78.8	453	35	US-10-940-774-7896
38	1443	78.2	390	10	US-08-663-617-35
39	1443	78.2	390	12	US-08-837-599-35
40	1443	78.2	390	22	US-09-791-537-117257
41	1425.5	77.2	386	22	US-09-791-537-114288
42	1425.5	77.2	386	30	US-10-403-571-76
43	1425.5	77.2	386	30	US-10-405-887-76
44	1421	77.0	386	22	US-09-791-537-36076
45	1349.5	73.1	399	22	US-09-791-537-1129

ALIGNMENTS

RESULT 1
US-09-787-491B-11
; Sequence 11, Application US/09787491B
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; TANG, Y. Tom;

Db 1 MetGlyAspProGluArgProGluAlaLaGlyLeuAspGlnAspGluArgSerSerSer 20
QY 61 GACACCAACGAAGTGAATAAGTCAATGAAGAGCCACTCTAAGAAGAGTTCTCGC 120
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
QY 121 CGGTTTGTCATCTTCCAAATCCAGTACCCTGATATTGGAAAATGATATAACAGGCACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCTCTGTCAGACAGCAGAGAGGTGACTTATCAAGAGTCTCCCTCAGTGAACAAG 240
Db 61 AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
QY 241 CTTAAAGCAGATGAGAAGTACTCTCTCATCTTACATCTTAGCTTTTTCAGCCAGCTGAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY 301 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTCTATGCGCTTCAAAATCTCATCGAGAATGTTCTCATCAGAGATGACAGTTTG 420
Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysLysAlaAspTyrAlaLeuArgTyrIleAlaAspArgLys 160
QY 481 ACCATGCCCTATGTTAAAGAAAAGCAGATTCGGGCTTCGGATGGATAGCAGATAGAAAA 540
Db 161 ThrMetProTyrValLysLysAlaAspTyrAlaLeuArgTyrIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTGTAAGAAGAGATTTTCTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGluValPhePheSerGly 200
QY 601 TCTTTTGTGCTATATTCTGGCTAAAGAAGAGAGGTCTTATGCCAGGACTCATTTTTC 660
Db 201 SerPheAlaAlaIlePheTyrLysLysAlaValGluGluValPhePheSerGly 220
QY 661 AATGAACATCATCAGCAGATGAGGACTTCACTGTGATGCTTGTGCTGATGTTCCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAAATTAAGCCTTCAGAAAGAGGTTCAGGAGATCATTTGATGCTGTCAAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
QY 781 ATTGACGAGGAGTTTAAACAGAGCCTTGCAGTGGCTCATTTGGAATGAATTGCATT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAAGTTCGATCTCA 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTCAGCAGAAAATCCTTTGATTTTATGAAAACATTCTTTAGAAGGAAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAAAATTTCTTTGAGAACAGCTTTTCAGAGTATCAGCGTTTTCGAGTTATGCGAGAAACC 1020
Db 321 ThrAsnPhePheGluLysArgValSerGlnTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGCTTTCACCTTGCATGCAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 3

US-10-170-205E-19045

; Sequence 19045, Application US/10170205E

; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
; TITLE OF INVENTION: CAPTURE AGENTS, AND USES THEREOF
; FILE REFERENCE: CL001381
; CURRENT APPLICATION NUMBER: US/10/170,205E
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19045
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-19045

Alignment Scores: 2.58e-186 Length: 351
Pred. No.: 1821.00 Matches: 351
Score: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.65% Indels: 0
DB: 27 Gaps: 0

US-10-698-228-12 (1-1053) x US-10-170-205E-19045 (1-351)

QY 1 ATCGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 MecGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAACGAAGTGAATAAGTCAAAATCAAAAGAGGCACCTCTTAAGAAAAGAGTTCTCGC 120
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
QY 121 CGGTTTGTCATCTTCCAAATCCAGTACCCTGATGATTTGGAAAATGATATAACAGGCACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCTCTGTCAGACAGCAGAGGTGACTTATCAAGAGATTCCTCAGTGAACAAG 240
Db 61 AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
QY 241 CTTAAAGCAGATGAGAAGTACTCTCATCTCAGAGTCTTAGCTTTTTCAGCCAGCTGAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY 301 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTCTATGCGCTTCAAAATCTCATCAGAGATGTTCTCATCAGAGATGACAGTTTG 420
Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAAGAAAAGCAGATTCGGGCTTCGGATGGATAGCAGATAGAAAA 540
Db 161 ThrMetProTyrValLysLysAlaAspTyrAlaLeuArgTyrIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTGTAAGAAGAGATTTTCTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGluValPhePheSerGly 200
QY 601 TCTTTTGTGCTATATTCTGGCTAAAGAAGAGAGGTCTTATGCCAGGACTCATTTTTC 660
Db 201 SerPheAlaAlaIlePheTyrLysLysAlaValGluGluValPhePheSerGly 220
QY 661 AATGAACATCATCAGCAGATGAGGACTTCACTGTGATGCTTGTGCTGATGTTCCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240


```

; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-990-328-7547

Alignment Scores:
Pred. No.: 2,58e-186 Length: 351
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservatative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.65% Indels: 0
DB: 35 Gaps: 0

US-10-698-228-12 (1-1053) x US-10-990-328-7547 (1-351)
QY 1 ATGGCGACCCGGAAGCGCGGCGCGCTGGATCAGGATCAGAGATCATCTTCA 60
DB 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAAGAAAGTCAAAATAAGTCAAAATGAAGAGCCACTCTCTAAAGAAAGAGTCTTCGC 120
DB 21 AspThrAsnGluSerGluLeuLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
QY 121 CGCTTTGTCTATCTTCCATCCAGTACCCCTGATATTGGAAATGTATTAACAGGCACAG 180
DB 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCTCTTCGACAGCAGAGAGGTGTGACTTATCAAGAGATCTCCCTCAGTGGAAACAAG 240
DB 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCATCTTAGCCCTTTTTCGAGCCAGCTGAT 300
DB 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY 301 GGAATTTGTAATAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGTTCAGAGGCT 360
DB 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTCTATGGCTTCCAAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTTCG 420
DB 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 480
DB 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTTAAGAAAAAAGCAGATTTGGCCCTTGGATGGATGAGATGAGAA 540
DB 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAGAGTGGGCTTGTCTGCTGTAGAGAGTCTTCTCTCAGCA 600
DB 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGCCTATATCTCGGCTTAAAGAGAGAGGTCTTATGCGAGACTCACTTTTTC 660
DB 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAATCATCAGCAGAGATGAAGCACTTCACTGTGACTTTGCTTGGCTGATGTTCCAA 720
DB 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAAATAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTCATGCTGTGCAA 780
DB 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
QY 781 ATTGAGCAGAGTTTTAAACAGAGGCTTGCAGAGTTCGCTCATTTGGAATGAATTCATT 840
DB 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTGTGGAATTCGATCTCA 900

```

```
Db      121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY      421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATCCAAATTGAA 480
Db      141 LeuIleAspThrTyrIleArgaspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY      481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGCCCTTCGATGGATAGCAGATAGAAAA 540
Db      161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaaspArgLys 180
QY      541 TCTACTTTGGGGAAGAGTGTGGCTTTGCTGCTGTGTAGAGAGATTCTTCTCTCAGGA 600
Db      181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY      601 TCTTTTGTGCTATATTCTGGCTTAAGAAAGAGAGGCTTATCCAGGACTCTTATCCGAGCTCTTTTCC 660
Db      201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY      661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGGCTGATGTTCCAA 720
Db      221 AsnGluLeuIleSerArgaspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY      721 TACTTAGTAATAAAGCTTTCAAGAAAGGCTCAGGAGATCATTTCTGATGCTGCAAA 780
Db      241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
QY      781 ATTGAGCAGGAGTTTACAGAGAGCTTCCAGAGCTTCCAGTGTGCTCATTTGGATGAATTCGATT 840
Db      261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY      841 TTGATGAACAGCTACATTGAGTTGTGAGCTGCAGAGATTACTTGTGCAACTTGGATTCTCA 900
Db      281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY      901 AAGGTTTTTCAGGCAGAAAAATCCTTTGATTTTATGAAAAACAATTTCTTTAGAGGAAAA 960
Db      301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY      961 ACAAAATTTCTTGAGAAACAGTTTCAGAGTATCAGCGTTTGCAGTTATGCGAGAAACC 1020
Db      321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY      1021 ACAGATAACGCTCTCACCTTGGATGCAGATTTT 1053
Db      341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
```

RESULT 7

```
US-60-449-629-94
; Sequence 94, Application US/60449629
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CANCER, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001449
; CURRENT APPLICATION NUMBER: US/60/449,629
; CURRENT FILING DATE: 2003-02-04
; NUMBER OF SEQ ID NOS: 2712
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 94
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-449-629-94
```

```
Alignment Scores:
Pred. No.: 2,58e-186 Length: 351
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.65% Indels: 0
DB: 37 Gaps: 0
```

```
US-10-698-228-12 (1-1053) x US-60-449-629-94 (1-351)
QY      1 ATGGCGACCCGCGAAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db      1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20
QY      61 GACACCAACGAAAGTGAATAAAGTCAAATCAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
Db      21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
QY      121 CGTTTGTGTCATCTTCCATCCAGTACCCTGATATTGGAAAAATGTATAACAGGCACAG 180
Db      41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
QY      181 GCITTCCTCTGGACAGCAGAGAGGTTGACTTATCAAGAGGATCTCCCTCACTGGAAACAAG 240
Db      61 AlaSerPheTrpThrAlaGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY      241 CTTTAAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTAGCCCTTTTTCAGCCCAAGTGA 300
Db      81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY      301 GGAATTTGTAATGAAATTTTGGTGGAGCGCTTTTAGTCAGAGGTCAGAGTTCAGAGGCT 360
Db      101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY      361 CGCTGTTTCTATCGCTTTCAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 420
Db      121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY      421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db      141 LeuIleAspThrTyrIleArgaspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY      481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGCCCTTCGATGGATAGCAGATAGAAAA 540
Db      161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY      541 TCTACTTTGGGGAAGAGTGTGGCTTTGCTGCTGTAGAGGAGTTTCTCTCAGGA 600
Db      181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY      601 TCTTTTGTGCTATATTCTGGCTTAAGAAAGAGAGGCTTATGCGCAGGACTCACTTTTTC 660
Db      201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY      661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTGCTGATGTTCCAA 720
Db      221 AsnGluLeuIleSerArgaspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY      721 TACTTAGTAATAAAGCTTTCAAGAAAGGCTCAGGAGATCATTTGATGCTGCTCAAA 780
Db      241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
QY      781 ATTGAGCAGGAGTTTAAACAGAGCCCTTGCAGTGTGGCCCTCATTTGGAATCAATTCGATT 840
Db      261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY      841 TTGATGAACAGTACATTGATGTTGTAGCTGCAGATTTACTTTGTGGAACCTTCGATTCTCA 900
Db      281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY      901 AAGTTTTTTCAGGCAGAAAAATCCTTTTGTATTTTATGAAAAACAATTTCTTTAGAGGAAAA 960
Db      301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY      961 ACAAAATTTCTTGAGAAACAGTTTTCAGAGTATCAGCGTTTGCAGTTATGCGAGAAACC 1020
Db      321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY      1021 ACAGATAACGCTCTCACCTTGGATGCAGATTTT 1053
Db      341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
```

RESULT 8
 US-60-505-218-322
 ; Sequence 322, Application US/60505218
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; FILE OF INVENTION: CANCER, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001482
 ; CURRENT APPLICATION NUMBER: US/60/505,218
 ; CURRENT FILING DATE: 2003-09-24
 ; NUMBER OF SEQ ID NOS: 22507
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 322
 ; LENGTH: 351
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-60-505-218-322

Alignment Scores:
 Pred. No.: 2 58e-186 Length: 351
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 98.65% Indels: 0
 DB: 37 Gaps: 0

US-10-698-228-12 (1-1053) x US-60-505-218-322 (1-351)

QY	1	ATGGGCGACCCGAAAGCGCGGAGCGGCGCTGGATCAGGATCAGAGATCATCTTCA	60
DB	1	MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer	20
QY	61	GACACCAAGAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAAGTTCCTCGC	120
DB	21	AspThrAsnGluSerGluIleIysSerAsnGluProLeuLeuArgIysSerSerArg	40
QY	121	CGGTTTGTCTATCTTCCATCCAGTACCTGATATTGGAAATGTATAACAGGCACAG	180
DB	41	ArgPheValIlePheProIleGlnTyrProAspIleTrpIysMetTyrLysGlnAla	60
QY	181	GCTTCCTCTCGACAGCAGAGAGGTGACTTATCAAGATCTCCTCACTGGAACAAG	240
DB	61	AlaSerPheTrpThrAlaGluValAspLeuSerIysAspLeuProHisTrpAsnLys	80
QY	241	CTTAAAGCAGATGAGAAGTACTTCACTCTCATCTTACATCTTAGCCCTTTTTCAGC	300
DB	81	LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAspAsp	100
QY	301	GGNATTTGAATGAATAATTTGTGGAGCGCTTTAGTCAGAGGTGAGGTCCAGAGGCT	360
DB	101	GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnProGluAla	120
QY	361	CGCTGTTTCTATGGCTTTTCAATTTCTATCGAGAATGTTCACTCAGAGATGTACAGT	420
DB	121	ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSer	140
QY	421	CTGATAGACACTTACATCAGATCCAGATCCCAAGAAAGGGAATTTTATTAATGCAAT	480
DB	141	LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu	160
QY	481	ACATGCCCTATGTTAAGAAAAAGCAGATTTGGCGCTTCGGATGGATAGCAGATAGAAA	540
DB	161	ThrMetProTyrValLysIysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys	180
QY	541	TCTACTTTTGGGAAGAGTGTGGCTTTGCTGTAGTAAGAGGAGTCTTCTCTCAGCA	600
DB	181	SerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePheSerGly	200
QY	601	TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGAGGTCTTATGCCAGACTCCTTTTC	660
DB	201	SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer	220

QY	661	AATGAACATCATCAGCAGAGATGAAGAGCTTCACTGTGACTTTGCTGCTGATGTTCCAA	720
DB	221	AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln	240
QY	721	TACTTAGTAATAGCTTTCAGAAAGAGGTACAGGAGATCATTTGATGCTGCTCTCAA	780
DB	241	TyrLeuValAsnLysProSerGluArgValArgGluIleIleValAspAlaValLys	260
QY	781	ATTGAGCAGGAGTTTTAAACAGAGCCTTGCCAGTTGGCCCTCATTTGGAATGAATTCAT	840
DB	261	IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle	280
QY	841	TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTTACTTTGGAACTTGGATTCTCA	900
DB	281	LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer	300
QY	901	AAGTTTTTTCAGCAGAAATCTTTTGGATTTTATGAAACATTTCTTTTGAAGAAAA	960
DB	301	LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys	320
QY	961	ACAAATTTCTTTCAGAAACGAGTTTCAGAGTATCAGCGTTTTCGAGTTTATGCGAGAAAC	1020
DB	321	ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr	340
QY	1021	ACAGATAACGCTTTCACCTTGGATGACAGATTTT 1053	
DB	341	ThrAspAsnValPheThrLeuAspAlaAspPhe 351	

RESULT 9
 US-09-791-537-27671
 ; Sequence 27671, Application US/09791537
 ; GENERAL INFORMATION:
 ; APPLICANT: Bionomix, Inc.
 ; APPLICANT: Debe, Derek
 ; APPLICANT: Danzer, Joseph
 ; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
 ; TITLE OF INVENTION: METHODS OF USE THEREOF
 ; FILE REFERENCE: 261/210
 ; CURRENT APPLICATION NUMBER: US/09/791,537
 ; CURRENT FILING DATE: 2001-02-22
 ; NUMBER OF SEQ ID NOS: 153055
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 27671
 ; LENGTH: 366
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-791-537-27671

Alignment Scores:
 Pred. No.: 2 62e-186 Length: 366
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 98.65% Indels: 0
 DB: 22 Gaps: 0

US-10-698-228-12 (1-1053) x US-09-791-537-27671 (1-366)

QY	1	ATGGGCGACCCGAAAGCGCGGAGCGGCGCTGGATCAGGATGAGAGATCATCTTCA	60
DB	16	MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer	35
QY	61	GACACCAAGAAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCCTCGC	120
DB	36	AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerSerArg	55
QY	121	CGGTTTGTCTATCTTCCATCCAGTACCTCGATATTGGAAATGTATAACAGGCACAG	180
DB	56	ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln	75
QY	181	GCTTCCTCTCGACAGCAGAGAGGTGACTTATCAAGGATCTCCCTCACTGGAACAAG	240
DB	76	AlaSerPheTrpThrAlaGluValAspLeuSerLysAspLeuProHisTrpAsnLys	95

```
QY 241 CTTAAAGCAGATGAGAAGTACTTCTCTCATCTTAGCCCTTTTTCAGCCAGTGAT 300
Db 96 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaSerAsp 115
QY 301 GGAATTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTCCAGAGCT 360
Db 116 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 135
QY 361 CGCTGTTCTATGGCTTCTCAATCTCATCGAGAATGTCACATCAGAGATGTACAGTTG 420
Db 136 ArgCysPheTyrGlyPheGlnIleuIleGluAsnValHisSerGluMetTyrSerLeu 155
QY 421 CTGATAGACACTTATCATCAGACATCCCAAGAAAGGAATTTTATTAATGCAATTGAA 480
Db 156 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 175
QY 481 ACCATGCCCTATGTTAAAGAAAAAGCAGATTTGGCGCTTCGATGATAGATAGAAAA 540
Db 176 ThrMetProTyrValLysLysLysAlaAspTTPAlaLeuArgTTPileAlaAspArgLys 195
QY 541 TCTACTTTTGGGGAAGAGTGTGGCTTTGCTGCTGTAGAGGATTTTCTCTCAGGA 600
Db 196 SerThrPheGlyGluArgValValAlaPheAlaValGluGlyValPhePheSerGly 215
QY 601 TCTTTTGTCTATATTCTGGCTTAAAGAGAGAGGTCTTATGCCAGACTCACTTTTTC 660
Db 216 SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 235
QY 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTTGTCTTGTCTGATGTTCCAA 720
Db 236 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 255
QY 721 TACTTAGTAAATAGCTTCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAAA 780
Db 256 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 275
QY 781 ATTGAGCAGAGTATTTTACAGAGCGCTTGCAGAGTGGCTCATTGGAATGAATTCATT 840
Db 276 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 295
QY 841 TTGATGAAACAGTACATTCAGTGTGTGAGCTGACAGATTAATCTGTGGAACCTTGGATTCTCA 900
Db 296 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 315
QY 901 AAGTTTTCAGCAGAGAAATCCTTTGATTTTATGGAATAAATCTTCTTTAGAGGAAAA 960
Db 316 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 335
QY 961 ACAAAATTTCTTTCAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC 1020
Db 336 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 355
QY 1021 ACAGATAACCTCTTCACTTGGATGATGATTTT 1053
Db 356 ThrAspAsnValPheThrLeuAspAlaAspPhe 366
```

RESULT 10

```
US-09-629-469A-12811
; Sequence 12811, Application US/09629469A
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: SAITO, KAORI
; APPLICANT: YAMAMOTO, JUNICHI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: WAKAMATSU, AI
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: OTSUKI, TETSUJI
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH CDNA AND THEIR USE
```

```
; FILE REFERENCE: 084335/0123
; CURRENT APPLICATION NUMBER: US/09/629,469A
; FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: JP 1999-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 1999-300253
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/183,322
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 19025
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12811
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-629-469A-12811
```

```
Alignment Scores:
Pred. No.: 6 97e-186 Length: 351
Score: 1817.00 Matches: 350
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.72% Mismatches: 0
Query Match: 98.43% Indels: 0
DB: 20 Gaps: 0
```

US-10-698-228-12 (1-1053) x US-09-629-469A-12811 (1-351)

```
QY 1 ATGGGCGACCCGGAAGCCGCGAGCGCGGTGATCAGATGAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAACGAAGTGAATAAGTCAAAATGAAGAGCCACTCCTTAGAAAAGTTCTGCG 120
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
QY 121 CGTTTGTCTACTTTCCTCAATCCAGTACCTGATATTTGGAAAAATCTATAAACAGCACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCTCTTCGACAGCAGAGAGGTGACTTATCAAGAGATCTCCCTCCTGGAACCAAG 240
Db 61 AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
QY 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCAGATCTTAGCCCTTTTTCGAGCCAGTGAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY 301 GGAATTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTCCAGAGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTCTATGGCTTCTCAATCTCATCAGAAATGTTCTACTCAGAGATGTACAGTTTG 420
Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACACTTATCATCAGAGATCCCAAGAAAAAGGGAATTTTATTAATGCAATTGAA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAAGAAAAAGCAGATTTGGCGCTTCGATGATAGATAGAAAA 540
Db 161 ThrMetProTyrValLysLysLysAlaAspTTPAlaLeuArgTTPileAlaAspArgLys 180
QY 541 TCTACTTTTGGGGAAGAGTGTGGCTTTGCTGCTGTAGAGGATTTTCTCTCAGGA 600
```


APPLICANT: ISHII, SHIZUKO
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: WAKAMATSU, AI
APPLICANT: NAGAI, KEIICHI
APPLICANT: OTSUKI, TETSUJI
TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE
FILE REFERENCE: 084335/0123
CURRENT APPLICATION NUMBER: US/10/917,503
CURRENT FILING DATE: 2004-08-13
PRIORITY FILING DATE: 2000-07-28
PRIORITY FILING DATE: 2000-07-28
PRIORITY FILING DATE: 1999-07-29
PRIORITY FILING DATE: 1999-08-27
PRIORITY FILING DATE: 2000-01-11
PRIORITY FILING DATE: 2000-05-02
PRIORITY FILING DATE: 2000-05-02
PRIORITY FILING DATE: 2000-06-09
PRIORITY FILING DATE: 2000-10-18
PRIORITY FILING DATE: 1999-10-18
PRIORITY FILING DATE: 2000-02-17
NUMBER OF SEQ ID NOS: 19025
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12811
LENGTH: 351
TYPE: PRT
ORGANISM: Homo sapiens
US-10-917-503-12811

Alignment Scores:
Pred. No.: 6,97e-186 Length: 351
Score: 1817.00 Matches: 350
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.72% Mismatches: 0
Query Match: 98.43% Indels: 0
DB: 35 Gaps: 0

US-10-698-228-12 (1-1053) x US-10-917-503-12811 (1-351)

QY 1 ATGGCGCCGCGAAGCGCGAGCGCGCGCTGATCAGATCAGATCATCTCA 60
DB 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAACGAAAGTCAATTAAGTCAATGAAGACCACTCTTAAGAAAGAGTCTCGC 120
DB 21 AspThrAsnGluSerGluLeuLeuSerAsnGluGluProLeuLeuArgLysSerArg 40
QY 121 CGTTTGTTCATCTTCCATCCAGTACCTGATATTTGGAAATGTTAAACAGGCACAG 180
DB 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCTCTTCGACGACGAGAGGTGACTTATCAAGATCTCCCTCACTGGAACAG 240
DB 61 AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
QY 241 CTTAAAGCAGATGAGAGTACTTCTATCTCATCTTACATCTTACCTTTTTCGAGCCAGTAT 300
DB 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGAGTGCGAGTTCAGAGGCT 360
DB 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTCTAGCGCTTTCAAATTTCTCATCGAGAATGTTCACTCAGAGATGACAGTTTG 420
DB 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACATTCATCATCAGATCCCAAGAAAGGAATTTTATTTAATGCAATGAA 480

Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCCCTTGGATGATGATGATGATGATGAT 540
Db 161 ThrMetProTyrValLysLysLysAlaAspIleAlaLeuArgTyrIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCCCTTCTGCTGCTAGAGAGGAGTTCCTTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaValGluGluValPhePheSerGly 200
QY 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGTCTTATGCCAGAGTCACTTTTTC 660
Db 201 SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAATCATCAGCAGAGATGAAGACTTTCACCTGTGCTGCTTGTCTGCTGATGTTCCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTTAATAAGCCTTTCAGAAAGAGGTTCAGGAGAGTCACTTGTGATGCTGCTCAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
QY 781 ATTGAGCAGAGTTCCTTAAACAGAGCCTTCCAGTTCGCTCATTCGATGATGATGATGAT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyValAsnCysIle 280
QY 841 TTGATGAACACAGTACATTTAGTTGTAGCTGACAGATTCCTGGAACCTTGGATTCCTCA 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTCAGCAGAAAAATCCTTTTGTGTTTATGTTTATGTTTATGTTTATGTTTATG 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAATTTCTTTCAGAAACAGTTCAGAGTATCAGCGTTCGAGTTCGAGTTCGAGTTCGAG 1020
Db 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACCTCTTCACCTTGGATGACAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 13

PCT-US02-10824-143
Sequence 143, Application PC/TUS0210824
GENERAL INFORMATION:
APPLICANT: Origene Technologies
TITLE OF INVENTION: Prostate Cancer Expression Profiles
FILE REFERENCE: 9U 206 PCT
CURRENT APPLICATION NUMBER: PCT/US02/10824
CURRENT FILING DATE: 2002-04-08
PRIOR APPLICATION NUMBER: US 60/281,732
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: US 60/281,731
PRIOR FILING DATE: 2001-04-06
NUMBER OF SEQ ID NOS: 211
SOFTWARE: PatentIn version 3.1
SEQ ID NO 143
LENGTH: 389
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US02-10824-143

Alignment Scores:
Pred. No.: 9.8e-147 Length: 389
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: 1 Gaps: 0

US-10-698-228-12 (1-1053) x PCT-US02-10824-143 (1-389)

```
QY 91 GAAGAGCCACTCTCTAAGAAAGAGTTCTCGCCGGTTTGTGTCATCTTTCAATCCAGTACCCT 150
Db 69 AspGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGluTyrHis 88
QY 151 GATATTGCAAAATGATATAACAGGCACAGCTTCTCTCGGACACAGAGAGTTGAC 210
Db 89 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 108
QY 211 TTATCAAGAGTCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAACTTCACTCTCT 270
Db 109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluArgTyrPheIleSer 128
QY 271 CACATCTTAGCTTTTTCAGCCAGTGTGTAATGTAATGAAATTTGGTGAGCGC 330
Db 129 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluLeuValGluArg 148
QY 331 TTATGTCAGAGGTGAGGTTCAGAGGCTCGCTGTTCTATGGCTTTCAATTCATC 390
Db 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
QY 391 GAGAGTGTTCACATCAGAGATGTACAGTTGCTGATAGACACTTACATCAGAGATCCCAAG 450
Db 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
QY 451 AAAAGGGAATTTTATTAAATGCAATGAAACCATGCCCTATCTTAAAGAAAAGACAGAT 510
Db 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysLysAlaAsp 208
QY 511 TGGGCTTTCGATGAGATGAGATAAATCTACTTTTGGGAAAAGAGTGTGGCTTT 570
Db 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaPhe 228
QY 571 GCTGCTGTAGAGGATTTCTCTCAGGATCTTTGCTGCTATATTCGGCTTAAAGAAG 630
Db 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 248
QY 631 AGAGGTCTTATGTCAGAGTCTACTTTTCCAATGAACTCATCAGCAGAGATGAAGACTT 690
Db 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
QY 691 CACTGTGACTTTGCTTGCCTGATGTTCCAAATACTTAGTAAATAAGCCTTCAGAAAGAG 750
Db 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
QY 751 GTCAGGAGATCATTTGCTGATGCTGCAAAATGAGCAGAGTGTTTTAAACAGAGCCTTG 810
Db 289 ValArgGluIleIleAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu 308
QY 811 CCAGTTGGCTCATTTGGAATGAACTTCATTTGATGAAACAGTACATTCAGTTGTAGCT 870
Db 309 ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla 328
QY 871 GACAGATTACTTTGGAATCTCGAATCTCAAAAGTTTTCAGCAGAAAATCTTTTGAT 930
Db 329 AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp 348
QY 931 TTTATGAAAAATTTCTTTAGAGAAAAACAAATTTCTTTGAGAAAACGAGTTTCAGAG 990
Db 349 PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu 368
QY 991 TATCAGCTTTTTCAGTTATGTCAGAAACACAGATAACGCTTTCACCTTGGATGCAGAT 1050
Db 369 TyrGlnArgMetGlyValMetSerProThrGluAsnSerPheThrLeuAspAlaAsp 388
QY 1051 TTT 1053
Db 389 Phe 389
```

RESULT 14

PCT-US03-17409-171

; Sequence 171, Application PC/TUS0317409

; GENERAL INFORMATION:

; APPLICANT: INCYTE CORPORATION

```
; APPLICANT: RICKERT, Paula K.
; APPLICANT: KRASNOW, Randi
; TITLE OF INVENTION: DIAGNOSTIC MARKERS FOR LUNG CANCER
; FILE REFERENCE: PA-0051 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/17409
; CURRENT FILING DATE: 2003-06-02
; PRIOR APPLICATION NUMBER: US 60/386,005
; PRIOR FILING DATE: 2002-06-04
; NUMBER OF SEQ ID NOS: 296
; SOFTWARE: PERL Program
; SEQ ID NO 171
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 000899CD1
PCT-US03-17409-171
```

Alignment Scores:

Pred. No.:	9.8e-147	Length:	389
Score:	1454.00	Matches:	269
Percent Similarity:	94.08%	Conservative:	33
Best Local Similarity:	83.80%	Mismatches:	19
Query Match:	78.76%	Indels:	0
DB:	1	Gaps:	0

US-10-698-228-12 (1-1053) x PCT-US03-17409-171 (1-389)

```
QY 91 GAAGAGCCACTCTCTAAGAAAGAGTTCTCGCCGGTTTGTGTCATCTTTCAATCCAGTACCCT 150
Db 69 AspGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGluTyrHis 88
QY 151 GATATTGCAAAATGATATAACAGGCACAGCTTCTCTCGGACACAGAGAGTTGAC 210
Db 89 AspIleTrpGlnMetTyrLysLysAlaGluAlaSerPheTrpThrAlaGluGluValAsp 108
QY 211 TTATCAAGAGTCTCCCTCACTGGAACAAGCTTAAAGCAGATGAGAACTTCACTCTCT 270
Db 109 LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluArgTyrPheIleSer 128
QY 271 CACATCTTAGCTTTTTCAGCCAGTGTGTAATGTAATGAAATTTGGTGAGCGC 330
Db 129 HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluLeuValGluArg 148
QY 331 TTATGTCAGAGGTGAGGTTCAGAGGCTCGCTGTTCTATGGCTTTCAATTCATC 390
Db 149 PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet 168
QY 391 GAGAGTGTTCACATCAGAGATGTACAGTTGCTGATAGACACTTACATCAGAGATCCCAAG 450
Db 169 GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys 188
QY 451 AAAAGGGAATTTTATTAAATGCAATGAAACCATGCCCTATCTTAAAGAAAAGACAGAT 510
Db 189 GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysLysAlaAsp 208
QY 511 TGGGCTTTCGATGAGATGAGATAAATCTACTTTTGGGAAAAGAGTGTGGCTTT 570
Db 209 TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaPhe 228
QY 571 GCTGCTGTAGAGGATTTCTCTCAGGATCTTTGCTGCTATATTCGGCTTAAAGAAG 630
Db 229 AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys 248
QY 631 AGAGGTCTTATGTCAGAGTCTACTTTTCCAATGAACTCATCAGCAGAGATGAAGACTT 690
Db 249 ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu 268
QY 691 CACTGTGACTTTGCTTGCCTGATGTTCCAAATACTTAGTAAATAAGCCTTCAGAAAGAG 750
Db 269 HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg 288
```


Qy 1051 TTT 1053
Db 389 Phe 389

Search completed: October 30, 2005, 06:48:03
Job time : 181.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 30, 2005, 06:28:55 ; Search time 36 Seconds
(without alignments)

13219.420 Million cell updates/sec

Title: US-10-698-228-12

Perfect score: 1846

Sequence: 1 atggcgacccggaaagcc.....tcaccttgatgcagatttt 1053

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 972358 seqs, 225972999 residues

Total number of hits satisfying chosen parameters: 1944716

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame_n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10698228/runat_26102005_100610_4069/app_query.fasta_1.2446
-DB=Pending Patents AA New -QPMT=fastan -SUFFIX=n2p.rapn -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10698228 -CGN 1_1_22 -runat_26102005_100610_4069
-NCPUS=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOC=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=130 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FPGAPOP=6 -FPGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending Patents AA New.*

1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US05_NEW_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US11_NEW_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1821	98.6	351	1	PCT-US05-09639-15
2	1821	98.6	351	1	PCT-US05-21650-34
3	1821	98.6	351	6	US-10-990-328A-7547
4	1821	98.6	351	7	US-11-088-686-15
5	1817	98.4	351	1	PCT-US05-09639-17
6	1817	98.4	351	7	US-11-088-686-17
7	1816	98.4	351	1	PCT-US05-09639-19
8	1816	98.4	351	7	US-11-088-686-19
9	1813	98.2	351	1	PCT-US05-09639-21
10	1813	98.2	351	7	US-11-088-686-21

11	1812	98.2	351	1	PCT-US05-09639-23	Sequence 23, Appl
12	1812	98.2	351	1	PCT-US05-09639-25	Sequence 25, Appl
13	1812	98.2	351	7	US-11-088-686-23	Sequence 23, Appl
14	1812	98.2	351	7	US-11-088-686-25	Sequence 25, Appl
15	1811	98.1	351	1	PCT-US05-09639-27	Sequence 27, Appl
16	1811	98.1	351	7	US-11-088-686-27	Sequence 27, Appl
17	1454	78.8	389	1	PCT-US03-10870-559	Sequence 559, App
18	1454	78.8	389	1	PCT-US05-09639-1	Sequence 1, Appl
19	1454	78.8	389	1	PCT-US05-21650-33	Sequence 33, Appl
20	1454	78.8	389	6	US-10-940-774A-6016	Sequence 6016, Ap
21	1454	78.8	389	7	US-11-088-686-1	Sequence 1, Appl
22	1454	78.8	449	6	US-10-990-328A-12349	Sequence 12349, A
23	1454	78.8	449	6	US-10-990-328A-12350	Sequence 12350, A
24	1454	78.8	453	6	US-10-450-763-45646	Sequence 45646, A
25	1454	78.8	453	6	US-10-940-774A-7896	Sequence 7896, Ap
26	1450	78.5	389	1	PCT-US05-09639-3	Sequence 3, Appl
27	1450	78.5	389	7	US-11-088-686-3	Sequence 3, Appl
28	1449	78.5	389	1	PCT-US05-09639-5	Sequence 5, Appl
29	1449	78.5	389	7	US-11-088-686-5	Sequence 5, Appl
30	1446	78.3	389	1	PCT-US05-09639-7	Sequence 7, Appl
31	1446	78.3	389	7	US-11-088-686-7	Sequence 7, Appl
32	1445	78.3	389	1	PCT-US05-09639-9	Sequence 9, Appl
33	1445	78.3	389	1	PCT-US05-09639-11	Sequence 11, Appl
34	1445	78.3	389	7	US-11-088-686-9	Sequence 9, Appl
35	1445	78.3	389	7	US-11-088-686-11	Sequence 11, Appl
36	1444	78.2	389	1	PCT-US05-09639-13	Sequence 13, Appl
37	1444	78.2	389	7	US-11-088-686-13	Sequence 13, Appl
38	1371	74.3	389	8	US-60-669-241-28508	Sequence 28508, A
39	1348	73.0	393	7	US-11-097-143-19182	Sequence 19182, A
40	1274	69.0	319	1	PCT-US05-12947-102	Sequence 102, App
41	1232.5	66.8	404	6	US-10-703-032-123588	Sequence 123588,
42	1156	62.6	330	7	US-11-251-208-421	Sequence 421, App
43	1147.5	62.2	329	7	US-11-251-208-427	Sequence 427, App
44	1128.5	61.1	327	7	US/11/241	GENERAL INFORMATI
45	1125.5	61.0	345	7	US-11-251-208-519	Sequence 519, App

ALIGNMENTS

RESULT 1

PCT-US05-09639-15
; Sequence 15, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004WOI
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-09639-15

Alignment Scores:
Pred. No.: 1.19e-180 Length: 351
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.65% Indels: 0
DB: 1 Gaps: 0

US-10-698-228-12 (1-1053) x PCT-US05-09639-15 (1-351)

QY 1 ATGGCGGACCCGGAAGCGCGGCGGTGGATGAGATGAGATCATCTTCA 60
|||||
Db 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSerSer 20
|||||
QY 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGACCACTCTTAAGAAAGATTCCTCGC 120

```
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
Qy 121 CGGTTTCTCATCTTCCAACTCCAGTACCTCATATTTGGAAATGATAAAGAGGACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
Qy 181 GCTTCTCTTCGGACAGAGAGAGTTCATCAAAAGATCTCCCTCACTGGAACAAG 240
Db 61 AlaSerPheTrpThrAlaGluValAspLysSerLysAspLeuProHisTrpAsnLys 80
Qy 241 CTTAAACAGATGAGAAGTACTTCTCATCTCTCACATCTTACGCTTTTTCAGCCAGTGAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Qy 301 GGAATGTTTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGCTCAGAGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGlnValGlnValProGluAla 120
Qy 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAATCTTCACTCAGAGATGTACAGTTTG 420
Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTGAA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Qy 481 ACCATGCCCTATGTTAAGAAAGAGAGATGGCGCTTTCAGAGGTGCAGGCTCAGAGCT 540
Db 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Qy 541 TCTACTTTTGGGAAAGAGTGGCTTTCAGAGGTCTTCTGCTAGAGAGTTCCTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
Qy 601 TCTTTTCTGCTATATTTCTGGTAAAGAGAGGTCTTATGCGAGAGTCTTCTTTC 660
Db 201 SerPheAlaAlaIlePheTrpLeuLysArgGlyLeuMetProGlyLeuThrPheSer 220
Qy 661 AATGAACCTATCAGCAGAGATGAAGGACTTCACTGTGCTTTCGCTGCTGATGTTCCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Qy 721 TACTTAGTAATAGCTTCTCAGAGAAAGGTCTCAGGAGATCATTTGTTGATGCTGTCAAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
Qy 781 ATTGAGCAGGAGTATTTTAAACAGAGCCTTGCAGTTCGCTCATTTGGAATGAATTCGATT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Qy 841 TTGATGAAACAGTACATGATGTTGTAGCTGACAGATTAATTGTGGAACTTGGATTCCTCA 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Qy 901 AAGTTTTCAGGACAGAAATCCTTTTGANTTTATGAAACATTTCTTTAGAGGAAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Qy 961 ACAATTTCTTTGAGAAACAGTTCAGAGTATCAGGCTTTTTCAGGTATGCGAGAAACC 1020
Db 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Qy 1021 ACAGATAACGCTTTCACCTTGGATGCGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
```

RESULT 2

```
PCT-US05-21650-34
; Sequence 34, Application PC/TUS0521650
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: MIGFS AS MODIFIERS OF THE IGF PATHWAY AND METHODS OF USE
```

```
; FILE REFERENCE: EX05-017
; CURRENT APPLICATION NUMBER: PCT/US05/21650
; CURRENT FILING DATE: 2005-06-27
; PRIOR APPLICATION NUMBER: 60/581,689
; PRIOR FILING DATE: 2004-06-21
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
; PCT-US05-21650-34

Alignment Scores:
Pred. No.: 1 19e-180 Length: 351
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.65% Indels: 0
DB: 1 Gaps: 0

US-10-698-228-12 (1-1053) x PCT-US05-21650-34 (1-351)
```

```
Qy 1 ATGGGCGACCCGGAAGCGCGGCTGGATCAGATGAGAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
Qy 61 GACACCAACCAAGTGAATTAAGTCAANTGAAGAGCCACTCCTAAGAAAGTTCGCG 120
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
Qy 121 CGGTTTCTCATCTTCCAAATCCAGTACCTCATATTTGGAAATGATAAAGAGGACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
Qy 181 GCTTCTCTTCGGACAGAGAGTTCATCTCTCACATCTTACAAAGGATCTCCCTCAGTGAACAAG 240
Db 61 AlaSerPheTrpThrAlaGluGluValAspLysSerLysAspLeuProHisTrpAsnLys 80
Qy 241 CTTAAACAGATGAGAAGTACTTCACTCTCACATCTTACGCTTTTTCGACGCGATGAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Qy 301 GGAATGTTTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGCTCAGAGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGlnValGlnValProGluAla 120
Qy 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAATCTTCACTCAGAGATGTACAGTTTG 420
Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTGAA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Qy 481 ACCATGCCCTATGTTAAGAAAGAGAGATGGCGCTTTCAGAGGTGCAGGCTCAGAGCT 540
Db 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Qy 541 TCTACTTTTGGGAAAGAGTGGCTTTCAGAGGTCTTCTGCTAGAGAGTTCCTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
Qy 601 TCTTTTCTGCTATATTTCTGGTAAAGAGAGGTCTTATGCGAGAGTCTTCTTTC 660
Db 201 SerPheAlaAlaIlePheTrpLeuLysArgGlyLeuMetProGlyLeuThrPheSer 220
Qy 661 AATGAACCTATCAGCAGAGATGAAGGACTTCACTGTGCTTTCGCTGCTGATGTTCCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Qy 721 TACTTAGTAATAGCTTCTCAGAGAAAGGTCTCAGGAGATCATTTGTTGATGCTGTCAAA 780
```

```
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluLeuValAlaValLys 260
QY 781 ATTGACGAGGTTTTTAACAGAGCCTTCCAGAGTGGCTCATTCGAATGATTCATT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAACAGTACATTGAGTTGTAGCTGACAGATTACTGTGGAACCTTGGATTCTCA 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTCCGACAGAAATCCTTTGATTTTATGCGAAACATCTTTAGAGGAAA 960
Db 301 LysValPheGlnAlaGlnAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAAAATTTCTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGCTTATGGCAGAAACC 1020
Db 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGTTCTTCACTTGGATGCGAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAlaAspPhe 351

RESULT 3
US-10-990-328A-7547
; Sequence 7547, Application US/10990328A
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328A
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7547
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-990-328A-7547

Alignment Scores:
Pred. No.: 1,19e-180 Length: 351
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.65% Indels: 0
DB: 6 Gaps: 0

US-10-698-228-12 (1-1053) x US-10-990-328A-7547 (1-351)
QY 1 ATGGGGACCCCGAAGCCGAGCGCGCTGGATCAGATGAGAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAAGCAAGTGAATGAATCAATGAAGCCACTCTTAAGAAGAGTTCTCC 120
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
QY 121 CGGTTTGTCTATCTTCCATCCAGTACCTCTGATATTGGAAAATGATATAACAGGCACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCCTTCTGACAGCAGAGAGGTGACTTATCAAGGATCTCCCTCACTGGAAACAAG 240
Db 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAGCAGATCAGAGTACTTCTCTCATCTTAGCCCTTTTTCAGCCAGTGTAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY 301 GGAATTTGTAATGAAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCGAGGTTCAGAGGCT 360
Db 301 GGAATTTGTAATGAAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCGAGGTTCAGAGGCT 360
```

```
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCAGAAATGTTCACTCAGAGATGTACAGTTTG 420
Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTATTAATGCAATTGAA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAGAAAAGACAGATGGCCCTTCCGATGCATGATGACAGATGAAA 540
Db 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAAGAGTGGCTTCTGCTCTGTAGAGAGGATTTTCTTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTCTGCTATATTCTGGCTAAAGAAGAGAGGTCTTATGCCAGGACTCACATTTTTC 660
Db 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAATCATCAGCAGAGATGAAGACTTCACCTGTGACTTTGCTGCTGATGTCCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAATTAAGCCTTCAGAGAAGAGGTGAGGAGAGATCATTTGATGCTGCTCAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
QY 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTGGCCCTCATTTGGAATGAATTCATTT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTTACTTGTGGAACCTTGGATTCTCA 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGTTTTTCAGGCAGAAAATCCTTTTGTATTTATGAAAACATTTCTTTAGAAGGAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAAAATTTCTTGAGAAACAGTTTTCAGAGTATCAGCGTTTTCAGCTTATGCGCAGAAACC 1020
Db 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGCTTCTTCACTTGGATGCGAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAlaAspPhe 351

RESULT 4
US-11-088-686-15
; Sequence 15, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-088-686-15

Alignment Scores:
Pred. No.: 1,19e-180 Length: 351
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Mismatches: 0
Conservative: 0
```

Best Local Similarity: 100.00%		Mismatches: 0
Query Match: 98.65%		Indels: 0
DB: 7		Gaps: 0
US-10-698-228-12 (1-1053) x US-11-088-686-15 (1-351)		
QY	1	ATGGGCGACCCGGAAGGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
DB	1	MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY	61	GACACCAAGAAAGTGAATAAGTCAATGAAGACCACTCTTAAGAAAGAGTTCTCGC 120
DB	21	AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerSerArg 40
QY	121	CGGTTTGTTCATCTTCCATCCAGTACCTGATATTTGGAATGTAATAACAGGCACAG 180
DB	41	ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
QY	181	GCTTCTCTCTGACAGCAGAGAGGTTGACTTATCAAGGATCTCCTCACTGGAACAAG 240
DB	61	AlaSerPheTyrThrAlaGluValAspLeuSerLysAspLeuProHisTyrPheAsnLys 80
QY	241	CTTAAAGCAGATGAGAGTACTTCTCATCTTACATCTTACATCTTACATCTTACATCTT 300
DB	81	LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaSerAsp 100
QY	301	GGAATTTGAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
DB	101	GlyIleValAsnGlnAsnLeuValGluArgPheSerGlnGlnValProGluAla 120
QY	361	CGCTGTTTCTATGGCTTTCAAAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
DB	121	ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 480
DB	141	LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY	481	ACATGCGCTATGTTAAGAAAAGCAGATTTGGCGCTTGGATGGATGAGATAGAAAA 540
DB	161	ThrMetProTyrValLysLysAlaAspTyrAlaLeuArgTyrPheAlaAspArgLys 180
QY	541	TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGTGTAAGAGAGTTTCTTCTCAGGA 600
DB	181	SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY	601	TCTTTTGTGCTATATCTGGCTTAAAGAGAGAGGTTCTTATGCCAGGACTCACTTTTTC 660
DB	201	SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY	661	AATGAACCTCATCAGCAGATGAGGACTTCTACTGTGACTTTGCTTGGCTGCTGATGTC 720
DB	221	AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY	721	TACTTAGTAAATAAGCCTTCAGAAGAAAGGTCAGGAGATCATCTTGTATGCTGTCAAA 780
DB	241	TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
QY	781	ATTGAGCAGGAGTTTAAACAGAGGCTTGCAGAGTTGGCTCATTTGGAATGAATTCATT 840
DB	261	IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY	841	TTGATGAAACAGTACATTGAGTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
DB	281	LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY	901	AGGTTTTTTCAGCAGAAAATCTTTTGTATTTATCGAAAACATTCTTTAGAGAAAA 960
DB	301	LysValPheGlnAlaGlnAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY	961	ACAAATTTCTTTCAGAAAACAGTTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020
DB	321	ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340

QY	1021	ACAGATAACGCTCTTTCACCTTGGATGAGATTTT 1053
DB	341	ThrAspAsnValPheThrLeuAspAlaAspPhe 351
RESULT 5		
PCT-US05-09639-17		
; Sequence 17, Application PC/TUS0509639		
; GENERAL INFORMATION:		
; APPLICANT: Yen, Yun		
; TITLE OF INVENTION: DRUG SCREENING		
; FILE REFERENCE: 14037-004W01		
; CURRENT APPLICATION NUMBER: PCT/US05/09639		
; CURRENT FILING DATE: 2005-03-23		
; PRIOR APPLICATION NUMBER: US 60/556,836		
; NUMBER OF SEQ ID NOS: 59		
; SOFTWARE: FastSeq for Windows Version 4.0		
; SEQ ID NO 17		
; LENGTH: 351		
; TYPE: PRT		
; ORGANISM: Homo sapiens		
PCT-US05-09639-17		
Alignment Scores:		
Pred. No.:	3 11e-180	Length: 351
Score:	1817.00	Matches: 350
Percent Similarity:	100.00%	Conservative: 1
Best Local Similarity:	99.72%	Mismatches: 0
Query Match:	98.43%	Indels: 0
DB:	1	Gaps: 0
US-10-698-228-12 (1-1053) x PCT-US05-09639-17 (1-351)		
QY	1	ATGGGCGACCCGGAAGGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
DB	1	MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY	61	GACACCAAGAAAGTGAATAAGTCAATGAAGACCACTCTTAAGAAAGAGTTCTCGC 120
DB	21	AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerSerArg 40
QY	121	CGGTTTGTCTATCTTCCAATCCAGTACCTGATATTTGGAATGTAATAACAGGCACAG 180
DB	41	ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
QY	181	GCTTCTCTCTGACAGCAGAGAGGTTGACTTATCAAGGATCTCCTCACTGGAACAAG 240
DB	61	AlaSerPheTyrThrAlaGluValAspLeuSerLysAspLeuProHisTyrPheAsnLys 80
QY	241	CTTAAAGCAGATGAGAGTACTTCTCATCTTACATCTTACATCTTACATCTTACATCTT 300
DB	81	LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaSerAsp 100
QY	301	GGAATTTGAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
DB	101	GlyIleValAsnGlnAsnLeuValGluArgPheSerGlnGlnValProGluAla 120
QY	361	CGCTGTTTCTATGGCTTTCAAAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
DB	121	ArgCysPheGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 480
DB	141	LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY	481	ACATGCGCTATGTTAAGAAAAGCAGATTTGGCGCTTGGATGGATGAGATAGAAAA 540
DB	161	ThrMetProTyrValLysLysAlaAspTyrAlaLeuArgTyrPheAlaAspArgLys 180
QY	541	TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGTGTAAGAGAGTTTCTTCTCAGGA 600
DB	181	SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200


```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-09639-19

Alignment Scores:
Pred. No.:      3,95e-180      Length:      351
Score:          1816.00      Matches:      350
Percent Similarity: 100.00%      Conservative: 1
Best Local Similarity: 99.72%      Mismatches: 0
Query Match:      98.37%      Indels:      0
DB:              1          Gaps:      0

US-10-698-228-12 (1-1053) x PCT-US05-09639-19 (1-351)

QY 1 ATGGCGCAGCCCGAAGCGCGGAGCGCGGCTGGATCAGATGAGAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAACGAAAGTAAATCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerSerArg 40
QY 121 CGTTTGTTCATCTTCCATCCAGTACCCTGATATTTGGAATATGATAAAGAGGACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCCTCTTGACAGCAGAGAGGTTGACTTATCAAGGATCTCCCTCACTGGAACAAG 240
Db 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAGCAGATGAGAGTACTTCTCATCTCTCAGATCAGGATCAGATCATCTTCA 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAspArg 100
QY 301 GCAATTTGTAATGAAATTTGGTGAGCGCTTTAGTCAGAGGTGCAGGTCAGAGGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnProGluAla 120
QY 361 CGCTGTTTCTATGGCTTTCAAAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db 121 ArgCysPheTrpGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAGAAAGAGAGATGGGCTTCGATGGATGAGATAGAGAA 540
Db 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAAGAGGAGTTTCTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTGCTATATTCTGGCTAAAGAGAGAGGCTTTATGCCAGGACTCACTTTTTC 660
Db 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACATCAGCAGAGATGAGGACTTCACTGTGACTTTGCTTGGCTGATGTTCCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAAATAAGCTTTCAGAAAGAGGTCAGGAGATCATCTGTGATGCTGTGCAAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
QY 781 ATTGAGCAGGAGTTTAAACAGAGAGCTTGGCCAGTGGCTCATTTGGAATGCAATT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
```

```
QY 841 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGTTTTTCAGCAGCAAAATCCTTTTGTGATTTTATGAAAACATTTCTTTAGAAGGAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAAAATTTCTTTCAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTATTCGCGAGAAACC 1020
Db 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGCTTTCACCTTTGGATGAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 8
US-11-088-686-19
; Sequence 19, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-088-686-19

Alignment Scores:
Pred. No.:      3,95e-180      Length:      351
Score:          1816.00      Matches:      350
Percent Similarity: 100.00%      Conservative: 1
Best Local Similarity: 99.72%      Mismatches: 0
Query Match:      98.37%      Indels:      0
DB:              7          Gaps:      0

US-10-698-228-12 (1-1053) x US-11-088-686-19 (1-351)

QY 1 ATGGCGCAGCCCGAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAACGAAAGTAAATGAAATGAAAGTCAAAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerSerArg 40
QY 121 CGTTTGTTCATCTTCCATCCAGTACCCTGATATTTGGAATATGATAAAGAGGACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCCTCTTGACAGCAGAGAGGTTGACTTATCAAGGATCTCCCTCACTGGAACAAG 240
Db 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAGCAGATGAGAGTACTTCTCATCTCTCAGATCAGGATCAGATCATCTTCA 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAspArg 100
QY 301 GCAATTTGTAATGAAATTTGGTGAGCGCTTTAGTCAGAGGTGCAGGTCAGAGGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnProGluAla 120
QY 361 CGCTGTTTCTATGGCTTTCAAAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db 121 ArgCysPheTrpGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAGAAAGAGAGATGGGCTTCGATGGATGAGATAGAGAA 540
Db 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAAGAGGAGTTTCTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTGCTATATTCTGGCTAAAGAGAGAGGCTTTATGCCAGGACTCACTTTTTC 660
Db 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACATCAGCAGAGATGAGGACTTCACTGTGACTTTGCTTGGCTGATGTTCCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAAATAAGCTTTCAGAAAGAGGTCAGGAGATCATCTGTGATGCTGTGCAAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
QY 781 ATTGAGCAGGAGTTTAAACAGAGAGCTTGGCCAGTGGCTCATTTGGAATGCAATT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
```

```
QY 421 CTGATAGACATTACATCAGAGATCCAGAAAAGGGAATTTTATTATGCAATTGAA 480
Db 141 LeuileAspThrTyIleArgaspProLysLysArgGluPheLeuPheAsnAlaileGlu 160
QY 481 ACCATGCCCTATGTTAAGAAAACAGATTGGGCTTCGGATGGATACAGATAGAAA 540
Db 161 ThrMetProTyValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTGGGAAAGAGTGGCTTGTCTGTAGAAGAGTTCCTTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTCTATATCTGGCTAAAGACAGAGTCTTATGCCAGGACTCCTTTTCC 660
Db 201 SerPheAlaAlailePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACATCAGCAGAGATGAGGACTTCTACTGTGACTTTCCTGCTGCTGATTCCAA 720
Db 221 AsnGluLeuileSerArgaspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAAATAAGCCTTCAGAAAAGGTCAGGAGATCATTGCTGCTGTCAAA 780
Db 241 TyrLeuValAsnLysProSerGluArgValArgGluileleValAspAlaValLys 260
QY 781 ATTGACGAGGAGTTCCTTAAAGAGCCTTGGCCTTCAGTGGCTCATTGGAATGCAATT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuileGlyMetAsnCysIle 280
QY 841 TTGATGAACAGTACATTGAGTTGTAGCTGACAGATTAATTGTGGAATTGCAATCTCA 900
Db 281 LeuMetLysGlnTyIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTCAGGCAGAAAATCCTTTTGAATTTATGGAACATTTCTTAGAAGAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluLys 320
QY 961 ACAAAATTTCTTGAAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTCATGACAGAAACC 1020
Db 321 ThrAsnPheGluLysArgValSerGluTyGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGCTTCACCTTGGATCGAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
```

RESULT 9
PCT-US05-09639-21
; Sequence 21, Application PC/TUS0509639

; GENERAL INFORMATION:
; APPLICANT: Yen. Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004WO1
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-09639-21

Alignment Scores:
Pred. No.: 8,12e-180 Length: 351
Score: 1813.00 Matches: 349
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.43% Mismatches: 0
Query Match: 98.21% Indels: 0
DB: 1 Gaps: 0

US-10-698-228-12 (1-1053) x PCT-US05-09639-21 (1-351)

RESULT 10

```
QY 1 ATCGCGGACCCGAAAGGCCGAGCCGGCGGTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAACGAAAGTGAATAAGTCAATCAAGAGCCACTCCTTAAGAAAGTTCTCC 120
Db 21 AspThrAsnGluSerGluileLysSerAsnGluProLeuLeuArgLysSerSerArg 40
QY 121 CGGTTTGTCTATCTTCCAAATCCAGTACCTCATATTTGGAAAAATGTAATAACAGC 180
Db 41 ArgPheValilePheProIleGlnTyProAspIleTrpLysMetTyLysGlnAlaGln 60
QY 181 GCTTCTCTTCGACAGCAGAGAGGTTGACTTATCAAGAGGATCTCCCTCCTCAGTGA 240
Db 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAGCAGATGAGAAGTACTTCTCATCTCTACATCTTATAGCCTTTTTCGACCGCA 300
Db 81 LeuLysAlaAspGluLysTyPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY 301 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAG 360
Db 101 GlyileValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATGCTTCAAAATTTCTCATCGAGATGTTCTACTCAGAGATGTACAGT 420
Db 121 ArgCysPhePheGlyPheGlnIleLeuileGluAsnValHisSerGluMetPheSerLeu 140
QY 421 CTGATAGACATTACATCAGAGATCCCAAGAAAAGGGAATTTTATTATGCAATTGAA 480
Db 141 LeuileAspThrTyIleArgaspProLysLysArgGluPheLeuPheAsnAlaileGlu 160
QY 481 ACCATGCCCTATGTTAAGAAAACAGAGATGCGCCTTCGATCGATGACAGATAGAAA 540
Db 161 ThrMetProTyValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTGGGAAAGAGTGGGCTTTCCTCTCTGTAGAAGAGTTCCTTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGGTCCTTATGCCAGGACTCCTTTTTC 660
Db 201 SerPheAlaAlailePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACATCAGCAGAGATGAGGACTTCTACTGTGACTTTCCTGCTGCTGATGTTCAA 720
Db 221 AsnGluLeuileSerArgaspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAAATAAGCCTTCAGAAAAGGTCAGGAGATCATTGTTGATGCTGTCAAA 780
Db 241 TyrLeuValAsnLysProSerGluArgValArgGluileleValAspAlaValLys 260
QY 781 ATTGACGAGGAGTTCCTTAAAGAGCCTTCGACAGTTCGCTCATTTGGAATGCAATT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuileGlyMetAsnCysIle 280
QY 841 TTGATGAACAGTACATTGAGTTGTAGCTGACAGATTAATTGTGGAATCTGGAATCTCA 900
Db 281 LeuMetLysGlnTyIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTCAGGCAGAAAATCCTTTTGAATTTATGGAACATTTCTTAGAAGAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluLys 320
QY 961 ACAAAATTTCTTGAAAACGAGTTTCAGAGTATCAGGAGTTCAGGAGTTCAGTTCATG 1020
Db 321 ThrAsnPheGluLysArgValSerGluTyGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGCTTCACCTTGGATCGAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
```

```
US-11-088-686-21
; Sequence 21, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-088-686-21

Alignment Scores:
Pred. No.:      8.12e-180      Length:      351
Score:          1813.00      Matches:     349
Percent Similarity: 100.00%      Conservative: 2
Best Local Similarity: 99.43%      Mismatches: 0
Query Match:      98.21%      Indels:      0
DB:               7          Gaps:          0

US-10-698-228-12 (1-1053) x US-11-088-686-21 (1-351)
QY 1 ATGGCGGACCCGGAAGCGCGGCGGCTGGATCAGAGTCAGAGATCATCTTCA 60
DB 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAAGCAAGTGAATGAAGTCAATGAAGCCACTCTAAGAAAGTCTCGC 120
DB 21 AspThrAsnGluSerGluLeuLysSerAsnGluProLeuLeuArgLysSerArg 40
QY 121 CGTTTGTTCATCTTCCCAATCCAGTACCCTGATATTTGGAATGATATAACAGGCACAG 180
DB 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCCTCTTGACAGCAGAGAGTGTACTTATCAAGGATCTCCCTCACTGGACAAG 240
DB 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAGCAGATGACAGAGTACTTCTCATCTCATCTCTAGCCCTTTTTCACGCCAGTAT 300
DB 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY 301 GGAATGTAAATGAATAATTTGGTGGAGCGCTTGTAGTCAGAGGTGCAGGTTCACAGAGCT 360
DB 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATGGCTTTCAAATTCATCGAGAATGTTTCACTCAGAGATGTACAGTTTG 420
DB 121 ArgCysPhePheGlyPheGlnIleLeuIleGluAsnValHisSerGluMetPheSerLeu 140
QY 421 CTGATAGACACTTACATCGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTGAA 480
DB 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGCCCTTCGATGGATAGCAGATAGAAA 540
DB 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGGAAGAGTGTGGCTTTGTGCTGTAGAGAGGAGTTTCTTCTCAGGA 600
DB 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGluValPhePheSerGly 200
QY 601 TCTTTTGTGCTATATCTGGCTAAAGAGAGAGGTCTTATCCAGGACTCATCTTTTCC 660
DB 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
```

```
DB 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAATAAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGCTCAAA 780
DB 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
QY 781 ATTGACGAGGAGTTTTTAACAGAAAGCTTGCACAGTTGGCCTCATTTGGAATGAATTGCATT 840
DB 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAAACAGTACATTGAGTTGTAGCTGACAGATTACTTGTGGAACTTGGATTCTTCA 900
DB 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTCAGCGACAGAAATCCTTTTGTATTTTATGAAAACATTTCTTTAGAGGAAAA 960
DB 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020
DB 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGCTCTTCACCTTGGATGCAGATTTT 1053
DB 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 11
PCT-US05-09639-23
; Sequence 23, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-09639-23

Alignment Scores:
Pred. No.:      1.03e-179      Length:      351
Score:          1812.00      Matches:     349
Percent Similarity: 100.00%      Conservative: 2
Best Local Similarity: 99.43%      Mismatches: 0
Query Match:      98.16%      Indels:      0
DB:               1          Gaps:          0

US-10-698-228-12 (1-1053) x PCT-US05-09639-23 (1-351)
QY 1 ATGGCGGACCCGGAAGCGCGGCGGCTGGATCAGAGTCAGAGATCATCTTCA 60
DB 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAAGCAAGTGAATGAAGTCAATGAAGCCACTCTAAGAAAGTCTCGC 120
DB 21 AspThrAsnGluSerGluLeuLysSerAsnGluProLeuLeuArgLysSerArg 40
QY 121 CGTTTGTTCATCTTCCCAATCCAGTACCCTGATATTTGGAATGATATAACAGGCACAG 180
DB 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCCTCTTGACAGCAGAGAGTGTACTTATCAAGGATCTCCCTCACTGGACAAG 240
DB 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAGCAGATGACAGAGTACTTCTCATCTCATCTCTAGCCCTTTTTCACGCCAGTAT 300
```

```
Db      81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY      301 CGAAATTGTAAATGAAATTTGGTGGAGCCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 360
Db      101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY      361 CGCTGTTTCTATGCGCTTTCAAAATTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG 420
Db      121 ArgCysPhePheGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTrpSerLeu 140
QY      421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTATGCAATGTAA 480
Db      141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY      481 ACCATGCCCTATGTTAAGAAAGACAGATTGGCCCTTCGGATGGATAGCAGATAGAAA 540
Db      161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY      541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTAGAGAGGCTTTCTTCTCAGGA 600
Db      181 SerThrPheGlyGluArgValAlaPheAlaValGluGlyValPhePheSerGly 200
QY      601 TCTTTTGTCTATATTCGGCTTAAAGAGAGAGGTCTTATGCGAGACTCACTTTTTC 660
Db      201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY      661 ATGAACCTCATCAGAGATCAGAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db      221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY      721 TACTTAGTAAATAAGCCTTCAAGAAAGGCTCAGGAGATCATCTGTCATGCTGTCAAA 780
Db      241 TyrLeuValAsnLysProSerGluGluArgValArgLysIleValAspAlaValLys 260
QY      781 ATTGACGAGGAGTGTAAACAGAGCCTTGCAGTTGGCTCATTTGGAATGAATTCATT 840
Db      261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY      841 TTGATGAACAGTACATTGAGTTGTAGCTGACAGATTACTGTGGAACTTGGATTCTCA 900
Db      281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY      901 AAGGTTTTTCAGGCAGAAATCCTTTGATTATTTATGAAACATTTCTTTAGAGCAAA 960
Db      301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY      961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTTGCAGTTATGGCAGAAACC 1020
Db      321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY      1021 ACAGATAACGCTTTCACCTTGGATCGAGATTTT 1053
Db      341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
```

RESULT 12

```
PCT-US05-09639-25
; Sequence 25, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004WO1
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-09639-25
```

```
Alignment Scores:
Pred. No.: 1,03e-179 Length: 351
Score: 1812.00 Matches: 349
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.43% Mismatches: 0
Query Match: 98.16% Indels: 0
DB: 1 Gaps: 0
US-10-698-228-12 (1-1053) x PCT-US05-09639-25 (1-351)
```

```
QY      1 ATGGGCGACCGCGAAAGCGCGGAGCGCGCTGGATCAGGATGAGAGATCATCTTCA 60
Db      1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20
QY      61 GACACCAACCAAGTGAATTAAGTCAATCAAGAGCCACTCTTAAGAAAGGTTCTCGC 120
Db      21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
QY      121 CGGTTTCTCATCTTTCAAATCCAGTACCTCATATTTGAAAAATGTATAAACAGGCACAG 180
Db      41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
QY      181 GCTTCTCTCTGACAGCAGAGAGGTTTCACTTATCAAGAGATCTCCCTCAGTCGAAACAAG 240
Db      61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY      241 CTTAAGACAGATGAAAGTACTTTCATCTCTCACATCTTAGCCTTTTTCGACGACGATGAT 300
Db      81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY      301 GGAATTGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
Db      101 GlyIleValAsnGluAsnLeuValArgPheSerGlnGluValGlnValProGluAla 120
QY      361 CGCTGTTTCTATGCGCTTCAAATTTCTCATCAGAAATGTTTCACTCAGAGATGTACAGTTTG 420
Db      121 ArgCysPheTrpGlyPheGlnIleLeuIleGluAsnValHisSerGluMetPheSerLeu 140
QY      421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAAATGCAATGAA 480
Db      141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY      481 ACCATGCCCTATGTTAAGAAAGACAGATTGGCCCTTCGGATGGATAGCAGATAGAAA 540
Db      161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY      541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTAGAGAGGATTTTCTTCTCAGGA 600
Db      181 SerThrPheGlyGluArgValAlaPheAlaValGluGlyValPhePheSerGly 200
QY      601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGTCTTATGCGAGGACTCACATTTTTC 660
Db      201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY      661 ATGAACCTCATCAGAGATGAAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db      221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY      721 TACTTAGTAAATAAGCCTTCAAGAAAGGCTCAGGAGATCATTTGTTGATGCTCTCAAA 780
Db      241 TyrLeuValAsnLysProSerGluGluArgValArgLysIleValAspAlaValLys 260
QY      781 ATTGACGAGGAGTGTAAACAGAGCCTTGCAGTTGGCTCATTTGGAATGAATGCAAT 840
Db      261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY      841 TTGATGAACAGTACATTGAGTTGTAGCTGACAGATTACTGTGGAACTTGGATTCTCA 900
Db      281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY      901 AAGGTTTTTCAGGCAGAAATCCTTTGATTATTTATGAAACATTTCTTTAGAGCAAA 960
```

```
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAATTTCTTTGAGAACGAGTTTCAGAGTATCAGCGTTTTCAGTATGTCAGGAAACC 1020
Db 321 ThrAsnPheGluLysArgValSerGluGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATACGCTCTTCACCTTGGATCGAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 13
US-11-088-686-23
; Sequence 23, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-088-686-23

Alignment Scores:
Pred. No.: 1,03e-179 Length: 351
Score: 1812.00 Matches: 349
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.43% Mismatches: 0
Query Match: 98.16% Indels: 0
DB: 7 Gaps: 0

US-10-698-228-12 (1-1053) x US-11-088-686-23 (1-351)
QY 1 ATGGCGGACCCGGAAGCGCGGCGGTCGATCAGATGAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAAGGAAAGTAAAGTCAATGAAGACCACTCTAAGAAAGAGTCTCCG 120
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
QY 121 CGTTTGTTCATCTTCCCAATCCAGTACCTGTATTTGAAAATGTATAAAGGACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCCTTCTGACAGCAGAGAGGTGACTATCAAGGATCTCCCTCACTGGAAACAAG 240
Db 61 AlaSerPheTyrTrpAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
QY 241 CTTAAGCAGATGAGAGTACTCTCTCATCTTAGCCTTTTTCAGCCAGTGAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAsp 100
QY 301 GGAATTTGAAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATGGCTTTCAAAATTCATPCGAGAATTTGTCACGAGATGTCAGTTTG 420
Db 121 ArgCysPhePheGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACACTTACATCAGATCCCAAGAAAAGGAAATTTTATTTAATGCAATGAA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATTGGCGCTTCGATGGATAGCAGATAGAAA 540
```

```
Db 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAAGAGTGGCTTGTCTGTCTAGAGGAGTTTCTTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTCTATATTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCCTTTTTC 660
Db 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACTCATCAGCAGAGATGAAGACTTCCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 221 AsnGluIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAAATAAGCTTCAGAGAAAGGCTCAGGAGAGATCATTTGTTGCTGCTCAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
QY 781 ATTGACGAGGAGTTTAAACAGAAAGCTTGCAGATTGGCCCTCATTTGGAATGAATTC 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAAACAGTACATTGAGTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTCAGGCAGAAAATCCTTTTGTATTTTATGAAACATTTCTTTAGAGGAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAATTTCTTTCAGAAACAGTTTTCAGAGTATCAGCGTTTTCAGGTTTATGCGAGAAACC 1020
Db 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGCTCTTCACCTTGGATCGAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 14
US-11-088-686-25
; Sequence 25, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-088-686-25

Alignment Scores:
Pred. No.: 1,03e-179 Length: 351
Score: 1812.00 Matches: 349
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.43% Mismatches: 0
Query Match: 98.16% Indels: 0
DB: 7 Gaps: 0

US-10-698-228-12 (1-1053) x US-11-088-686-25 (1-351)
QY 1 ATGGCGGACCCGGAAGCGCGGCGGTCGATCAGATGAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAAGGAAAGTAAAGTCAATGAAGACCACTCTAAGAAAGAGTCTCCG 120
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
QY 121 CGTTTGTTCATCTTCCCAATCCAGTACCTGTATTTGAAAATGTATAAAGGACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCCTTCTGACAGCAGAGAGGTGACTATCAAGGATCTCCCTCACTGGAAACAAG 240
Db 61 AlaSerPheTyrTrpAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
QY 241 CTTAAGCAGATGAGAGTACTCTCTCATCTTAGCCTTTTTCAGCCAGTGAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAsp 100
QY 301 GGAATTTGAAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATGGCTTTCAAAATTCATPCGAGAATTTGTCACGAGATGTCAGTTTG 420
Db 121 ArgCysPhePheGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACACTTACATCAGATCCCAAGAAAAGGAAATTTTATTTAATGCAATGAA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATTGGCGCTTCGATGGATAGCAGATAGAAA 120
```

```
Db 21 AspThrAsnGluSerGluLeuIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
QY 121 CGGTTTGTCACTTTCCAAATCCAGTACCTCATATTTTGGAAATGTATAAAGCAGGCACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCCTTCTGGACAGCAGAGAGGTTGACTTATCAAGGATCTCCCTCAGTGGAAACAG 240
Db 61 AlaSerPheTyrTrpAlaGluGluValAspLeuSerLysAspLeuProHisTyrPheAsnLys 80
QY 241 CTTAAGCAGATGAGAGTACTTCACTCTCATCTTACATCTTAGCCCTTTTTCGACCCAGTGT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY 301 GGAATTTGAAATGAAATTTGGTGAGCGCTTACTAGAGGAGTGCAGGCTTCCAGAGGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAATGTTTCACTCAGAGATGTACAGTTTG 420
Db 121 ArgCysPheTyrPheGlnIleLeuIleGluAsnValHisSerGluMetPheSerLeu 140
QY 421 CTGATAGACATTCATCAGATCCCAAGAAAGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTAAGAAAAAGCAGATGGGCTTCCGATGATACGATAGCAATGAAAA 540
Db 161 ThrMetProTyrValLysLysLysAlaAspTyrAlaLeuArgTyrPheAlaAspArgLys 180
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTGTGCTGTAGAGAGGTTTCTTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValAlaGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTCTATATTCTGGCTAAAGAGAGGCTTTATGCCAGGACTCACTTTTTC 660
Db 201 SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACTCATCAGAGATGAGGACTTCACTGTGACTTTGCTGCTGATGTTCCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAAATTAAGCTTTCAGAAAGAGGCTCAGGAGATCATTTGTGCTGTCAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleValAspAlaValLys 260
QY 781 ATTGACGAGGAGTTTAAACAGAGCCCTTGCAGTGGCCTCATTTGGAATGCAATT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAAACAGTACATGAGTTGTAGCTCAGAGATTAATTTGGAACTTGGATCTCA 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTCAGGCAGAAAACTTTTTCATTTTATGGAACATTTCTTTAGAGGAAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAAAATTTCTTTGAGAACAGGTTTTCAGAGTATCAGCGTTTGTGAGTATGCGAAAAACC 1020
Db 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATACGTTCTTCACTTGGATGCGAGTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
```

RESULT 15

```
PCT-US05-09639-27
; Sequence 27, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004WO1
```

```
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-09639-27
```

Alignment Scores:

```
Pred. No.: 1,31e-179 Length: 351
Score: 1811.00 Matches: 349
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.43% Mismatches: 0
Query Match: 98.10% Indels: 0
DB: 1 Gaps: 0
```

US-10-698-228-12 (1-1053) x PCT-US05-09639-27 (1-351)

```
QY 1 ATGGGCGACCCGAAAGCGGAGCGGCGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAACGAAAGTCAAAATTAAGTCAATGAAGAGCCACTCCTAAGAAAGAGTCTCGC 120
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
QY 121 CGGTTTGTCTATCTTTCCTCAATCCAGTACCTGATATTGGAAATGTATAAAGCAGCAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCTCTCTGGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCAGTGGAAACAG 240
Db 61 AlaSerPheTyrTrpAlaGluGluValAspLeuSerLysAspLeuProHisTyrPheAsnLys 80
QY 241 CTTAAGCAGATGAGAGTACTTCACTCTCATCTTACATCTTAGCCCTTTTTCGACCCAGTGT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY 301 GGAATTTGAAATGAAATTTGGTGAGCGCTTACTAGAGGAGTGCAGGTTCCAGAGGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATGGCTTTCAAAATTCATCGAGAATGTTTCACTCAGAGATGTACAGTTTG 420
Db 121 ArgCysPheTyrPheGlnIleLeuIleGluAsnValHisSerGluMetTyrPheSerLeu 140
QY 421 CTGATAGACATTCATCAGATCCCAAGAAAGCGGCTGGATCAGGATGAGAGATCATCTTCA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTAAGAAAAAGCAGATGGGCTTCCGATGATACGATAGCAATGAAAA 540
Db 161 ThrMetProTyrValLysLysLysAlaAspTyrAlaLeuArgTyrPheAlaAspArgLys 180
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTGTGCTGTAGAGAGGTTTCTTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValAlaGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTCTATATTCTGGCTAAAGAGAGGCTTTATGCCAGGACTCACTTTTTC 660
Db 201 SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACTCATCAGAGATGAGGACTTCACTGTGACTTTGCTGCTGATGTTCCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAAATTAAGCTTTCAGAAAGAGGCTCAGGAGATCATTTGTGCTGTCAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleValAspAlaValLys 260
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2005, 23:48:01 ; Search time 3214.5 Seconds
(without alignments)
13372.738 Million cell updates/sec

Title: US-10-698-228-12
Perfect score: 1053
Sequence: 1 atggggaccgcgaaggcc.....tcaccttgatgcagatttt 1053

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 45554873 seqs, 20411521753 residues
Total number of hits satisfying chosen parameters: 91109746

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents NA Main:
1: /cgn2_6/ptodata/1/pna/PCTUS1 COMB.seq:
2: /cgn2_6/ptodata/1/pna/PCTUS2 COMB.seq:
3: /cgn2_6/ptodata/1/pna/PCTUS COMB.seq:
4: /cgn2_6/ptodata/1/pna/US06 COMB.seq:
5: /cgn2_6/ptodata/1/pna/US07 COMB.seq:
6: /cgn2_6/ptodata/1/pna/US08 COMB.seq:
7: /cgn2_6/ptodata/1/pna/US081 COMB.seq:
8: /cgn2_6/ptodata/1/pna/US082 COMB.seq:
9: /cgn2_6/ptodata/1/pna/US083 COMB.seq:
10: /cgn2_6/ptodata/1/pna/US084 COMB.seq:
11: /cgn2_6/ptodata/1/pna/US085 COMB.seq:
12: /cgn2_6/ptodata/1/pna/US086 COMB.seq:
13: /cgn2_6/ptodata/1/pna/US087 COMB.seq:
14: /cgn2_6/ptodata/1/pna/US088 COMB.seq:
15: /cgn2_6/ptodata/1/pna/US089 COMB.seq:
16: /cgn2_6/ptodata/1/pna/US090 COMB.seq:
17: /cgn2_6/ptodata/1/pna/US091 COMB.seq:
18: /cgn2_6/ptodata/1/pna/US092A COMB.seq:
19: /cgn2_6/ptodata/1/pna/US092B COMB.seq:
20: /cgn2_6/ptodata/1/pna/US093A COMB.seq:
21: /cgn2_6/ptodata/1/pna/US093B COMB.seq:
22: /cgn2_6/ptodata/1/pna/US094 COMB.seq:
23: /cgn2_6/ptodata/1/pna/US095A COMB.seq:
24: /cgn2_6/ptodata/1/pna/US095B COMB.seq:
25: /cgn2_6/ptodata/1/pna/US095C COMB.seq:
26: /cgn2_6/ptodata/1/pna/US095D COMB.seq:
27: /cgn2_6/ptodata/1/pna/US096A COMB.seq:
28: /cgn2_6/ptodata/1/pna/US096B COMB.seq:
29: /cgn2_6/ptodata/1/pna/US096C COMB.seq:
30: /cgn2_6/ptodata/1/pna/US096D COMB.seq:
31: /cgn2_6/ptodata/1/pna/US096E COMB.seq:
32: /cgn2_6/ptodata/1/pna/US097A COMB.seq:
33: /cgn2_6/ptodata/1/pna/US097B COMB.seq:
34: /cgn2_6/ptodata/1/pna/US097C COMB.seq:
35: /cgn2_6/ptodata/1/pna/US098A COMB.seq:
36: /cgn2_6/ptodata/1/pna/US098B COMB.seq:
37: /cgn2_6/ptodata/1/pna/US098C COMB.seq:
38: /cgn2_6/ptodata/1/pna/US098D COMB.seq:
39: /cgn2_6/ptodata/1/pna/US099A COMB.seq:
40: /cgn2_6/ptodata/1/pna/US099B COMB.seq:
41: /cgn2_6/ptodata/1/pna/US099C COMB.seq:
42: /cgn2_6/ptodata/1/pna/US099D COMB.seq:
43: /cgn2_6/ptodata/1/pna/US099E COMB.seq:

44: /cgn2_6/ptodata/1/pna/US099F COMB.seq:
45: /cgn2_6/ptodata/1/pna/US099G COMB.seq:
46: /cgn2_6/ptodata/1/pna/US100A COMB.seq:
47: /cgn2_6/ptodata/1/pna/US100B COMB.seq:
48: /cgn2_6/ptodata/1/pna/US101A COMB.seq:
49: /cgn2_6/ptodata/1/pna/US101B COMB.seq:
50: /cgn2_6/ptodata/1/pna/US102A COMB.seq:
51: /cgn2_6/ptodata/1/pna/US102B COMB.seq:
52: /cgn2_6/ptodata/1/pna/US103A COMB.seq:
53: /cgn2_6/ptodata/1/pna/US103B COMB.seq:
54: /cgn2_6/ptodata/1/pna/US104A COMB.seq:
55: /cgn2_6/ptodata/1/pna/US104B COMB.seq:
56: /cgn2_6/ptodata/1/pna/US105A COMB.seq:
57: /cgn2_6/ptodata/1/pna/US105B COMB.seq:
58: /cgn2_6/ptodata/1/pna/US106A COMB.seq:
59: /cgn2_6/ptodata/1/pna/US107A COMB.seq:
60: /cgn2_6/ptodata/1/pna/US107B COMB.seq:
61: /cgn2_6/ptodata/1/pna/US107C COMB.seq:
62: /cgn2_6/ptodata/1/pna/US107D COMB.seq:
63: /cgn2_6/ptodata/1/pna/US108A COMB.seq:
64: /cgn2_6/ptodata/1/pna/US108B COMB.seq:
65: /cgn2_6/ptodata/1/pna/US109A COMB.seq:
66: /cgn2_6/ptodata/1/pna/US109B COMB.seq:
67: /cgn2_6/ptodata/1/pna/US109C COMB.seq:
68: /cgn2_6/ptodata/1/pna/US110 COMB.seq:
69: /cgn2_6/ptodata/1/pna/US6000 COMB.seq:
70: /cgn2_6/ptodata/1/pna/US6001 COMB.seq:
71: /cgn2_6/ptodata/1/pna/US6002 COMB.seq:
72: /cgn2_6/ptodata/1/pna/US6003 COMB.seq:
73: /cgn2_6/ptodata/1/pna/US6004 COMB.seq:
74: /cgn2_6/ptodata/1/pna/US6005 COMB.seq:
75: /cgn2_6/ptodata/1/pna/US6006 COMB.seq:
76: /cgn2_6/ptodata/1/pna/US6007 COMB.seq:
77: /cgn2_6/ptodata/1/pna/US6008 COMB.seq:
78: /cgn2_6/ptodata/1/pna/US6009 COMB.seq:
79: /cgn2_6/ptodata/1/pna/US6010 COMB.seq:
80: /cgn2_6/ptodata/1/pna/US6011 COMB.seq:
81: /cgn2_6/ptodata/1/pna/US6012 COMB.seq:
82: /cgn2_6/ptodata/1/pna/US6013 COMB.seq:
83: /cgn2_6/ptodata/1/pna/US6014 COMB.seq:
84: /cgn2_6/ptodata/1/pna/US6015 COMB.seq:
85: /cgn2_6/ptodata/1/pna/US6016 COMB.seq:
86: /cgn2_6/ptodata/1/pna/US6017 COMB.seq:
87: /cgn2_6/ptodata/1/pna/US6018 COMB.seq:
88: /cgn2_6/ptodata/1/pna/US6019 COMB.seq:
89: /cgn2_6/ptodata/1/pna/US6020 COMB.seq:
90: /cgn2_6/ptodata/1/pna/US6021 COMB.seq:
91: /cgn2_6/ptodata/1/pna/US6022 COMB.seq:
92: /cgn2_6/ptodata/1/pna/US6023A COMB.seq:
93: /cgn2_6/ptodata/1/pna/US6023B COMB.seq:
94: /cgn2_6/ptodata/1/pna/US6024 COMB.seq:
95: /cgn2_6/ptodata/1/pna/US6025 COMB.seq:
96: /cgn2_6/ptodata/1/pna/US6026 COMB.seq:
97: /cgn2_6/ptodata/1/pna/US6027 COMB.seq:
98: /cgn2_6/ptodata/1/pna/US6028 COMB.seq:
99: /cgn2_6/ptodata/1/pna/US6029 COMB.seq:
100: /cgn2_6/ptodata/1/pna/US6030 COMB.seq:
101: /cgn2_6/ptodata/1/pna/US6031 COMB.seq:
102: /cgn2_6/ptodata/1/pna/US6032 COMB.seq:
103: /cgn2_6/ptodata/1/pna/US6033 COMB.seq:
104: /cgn2_6/ptodata/1/pna/US6034 COMB.seq:
105: /cgn2_6/ptodata/1/pna/US6035 COMB.seq:
106: /cgn2_6/ptodata/1/pna/US6036 COMB.seq:
107: /cgn2_6/ptodata/1/pna/US6037 COMB.seq:
108: /cgn2_6/ptodata/1/pna/US6038 COMB.seq:
109: /cgn2_6/ptodata/1/pna/US6039 COMB.seq:
110: /cgn2_6/ptodata/1/pna/US6040 COMB.seq:
111: /cgn2_6/ptodata/1/pna/US6041 COMB.seq:
112: /cgn2_6/ptodata/1/pna/US6042 COMB.seq:
113: /cgn2_6/ptodata/1/pna/US6043 COMB.seq:
114: /cgn2_6/ptodata/1/pna/US6044 COMB.seq:
115: /cgn2_6/ptodata/1/pna/US6045 COMB.seq:
116: /cgn2_6/ptodata/1/pna/US6046 COMB.seq:

117: /cgn2_6/ptodata/1/pna/US6047_COMB.seq.*
 118: /cgn2_6/ptodata/1/pna/US6048_COMB.seq.*
 119: /cgn2_6/ptodata/1/pna/US6049_COMB.seq.*
 120: /cgn2_6/ptodata/1/pna/US6050_COMB.seq.*
 121: /cgn2_6/ptodata/1/pna/US6051_COMB.seq.*
 122: /cgn2_6/ptodata/1/pna/US6052_COMB.seq.*
 123: /cgn2_6/ptodata/1/pna/US6053_COMB.seq.*
 124: /cgn2_6/ptodata/1/pna/US6054_COMB.seq.*
 125: /cgn2_6/ptodata/1/pna/US6055_COMB.seq.*
 126: /cgn2_6/ptodata/1/pna/US6056_COMB.seq.*
 127: /cgn2_6/ptodata/1/pna/US6057_COMB.seq.*
 128: /cgn2_6/ptodata/1/pna/US6058_COMB.seq.*
 129: /cgn2_6/ptodata/1/pna/US6059_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1053	100.0	1053	58	US-10-698-228-12	Sequence 12, Appl
2	1053	100.0	1081	58	US-10-698-228-4	Sequence 4, Appl
3	1051.4	99.8	1053	58	US-10-698-228-2	Sequence 2, Appl
4	1051.4	99.8	2596	34	US-09-787-491B-28	Sequence 28, Appl
5	1051.4	99.8	2596	81	US-60-128-660-20	Sequence 20, Appl
6	1051.4	99.8	3397	49	US-10-170-235-20856	Sequence 20856, A
7	1051.4	99.8	4955	1	PCT-US02-18947-71	Sequence 71, Appl
8	1051.4	99.8	4955	49	PCT-US04-01100-71	Sequence 71, Appl
9	1051.4	99.8	4955	49	US-10-172-118-71	Sequence 71, Appl
10	1051.4	99.8	4955	52	US-10-342-887-71	Sequence 71, Appl
11	1051.4	99.8	4955	58	US-10-698-228-3	Sequence 3, Appl
12	1051	99.8	1849	67	US-10-990-328-49	Sequence 49, Appl
13	1051	99.8	1849	120	US-60-505-218-20	Sequence 20, Appl
14	1051	99.8	3397	114	US-60-449-629-12	Sequence 12, Appl
15	1048.2	99.5	1601	28	US-09-629-469A-12810	Sequence 12810, A
16	1048.2	99.5	1601	65	US-10-917-503-12810	Sequence 12810, A
17	1040.4	98.8	2597	97	US-60-278-258-266	Sequence 266, App
18	980.4	93.1	5609	26	US-09-584-852-6398	Sequence 6398, App
19	878.4	83.4	1140	48	US-10-144-771-870	Sequence 870, App
20	878.4	83.4	1140	106	US-60-360-207-870	Sequence 870, App
21	850.6	61.8	706	2	PCT-US03-13853-8394	Sequence 8394, App
22	850.6	61.8	706	56	US-10-513-369-8394	Sequence 8394, App
23	850.6	61.8	706	107	US-60-377-240-8394	Sequence 8394, App
24	580.6	55.1	2500	29	US-09-658-659B-9	Sequence 9, Appl
25	580.6	55.1	2500	29	US-09-658-659B-9	Sequence 9, Appl
26	580	55.1	2482	1	PCT-US01-04926A-514	Sequence 514, App
27	580	55.1	2482	23	US-09-515-126-13666	Sequence 13666, A
28	580	55.1	2482	26	US-09-577-408-13666	Sequence 13666, A
29	580	55.1	2482	50	US-10-220-335-514	Sequence 514, App
30	580	55.1	2482	51	US-10-290-752-514	Sequence 514, App
31	580	55.1	2500	1	PCT-US02-10824-4	Sequence 4, Appl
32	580	55.1	2500	42	US-09-954-456-724	Sequence 724, App
33	580	55.1	2500	42	US-09-954-456-724	Sequence 1169, App
34	580	55.1	2500	42	US-09-954-456-1827	Sequence 1827, App
35	580	55.1	2500	58	US-10-641-643-1370	Sequence 1370, App
36	580	55.1	2500	62	US-10-733-878-458	Sequence 458, App
37	580	55.1	2500	62	US-10-756-149-713	Sequence 713, App
38	580	55.1	2500	63	US-10-843-641A-3751	Sequence 3751, App
39	580	55.1	2500	63	US-10-843-641A-4196	Sequence 4196, App
40	580	55.1	2500	63	US-10-843-641A-4854	Sequence 4854, App
41	580	55.1	2500	66	US-10-940-774-145	Sequence 145, App
42	580	55.1	2500	68	US-11-000-688-1211	Sequence 1211, App
43	580	55.1	2500	126	US-60-568-073-328	Sequence 328, App
44	578.4	54.9	1649	2	PCT-US04-09289-583	Sequence 583, App
45	578.4	54.9	1989	1	PCT-US00-05882-505	Sequence 505, App

ALIGNMENTS

RESULT 1
 US-10-698-228-12
 ; Sequence 12, Application US/10698228
 ; GENERAL INFORMATION:
 ; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
 ; TITLE OF INVENTION: New Protein and its DNA
 ; FILE REFERENCE: 2619WOOP
 ; CURRENT APPLICATION NUMBER: US/10/698,228
 ; CURRENT FILING DATE: 2003-10-30
 ; PRIOR APPLICATION NUMBER: US/10/019,733
 ; PRIOR FILING DATE: 2001-12-28
 ; PRIOR APPLICATION NUMBER: JP 11-181131
 ; PRIOR FILING DATE: 1999-06-28
 ; PRIOR APPLICATION NUMBER: JP 11-192391
 ; PRIOR FILING DATE: 1999-07-06
 ; PRIOR APPLICATION NUMBER: JP 2000-017770
 ; PRIOR FILING DATE: 2000-01-21
 ; NUMBER OF SEQ ID NOS: 14
 ; SEQ ID NO 12
 ; LENGTH: 1053
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION:
 US-10-698-228-12
 Query Match 100.0%; Score 1053; DB 58; Length 1053;
 Best local Similarity 100.0%; Pred. No. 6.1e-276;
 Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGGGGCACCGGAAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
 DB 1 ATGGGGCACCGGAAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
 QY 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGACCCACTCTTAAGAAAGAGTTCTGC 120
 DB 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGACCCACTCTTAAGAAAGAGTTCTGC 120
 QY 121 CGGTTTGTATCTTTCATTCAGATCCAGTACCTGTATTTGGAAATGTATAACAGGCACAG 180
 DB 121 CGGTTTGTATCTTTCATTCAGATCCAGTACCTGTATTTGGAAATGTATAACAGGCACAG 180
 QY 181 GCTTCTCTTGGACAGCAGAGAGGTTGACTTATCAAGAGGATCTCCCTCACTGGAAACAAG 240
 DB 181 GCTTCTCTTGGACAGCAGAGAGGTTGACTTATCAAGAGGATCTCCCTCACTGGAAACAAG 240
 QY 241 CTTAAAGCAGATGAGAAAGTACTTTCATCTCTCACATCTTTAGCCTTTTTTGCAGCCAGTGAT 300
 DB 241 CTTAAAGCAGATGAGAAAGTACTTTCATCTCTCACATCTTTAGCCTTTTTTGCAGCCAGTGAT 300
 QY 301 GGAATTGTAAATGAAATTTGGTGGAGCGCTTTTAGTCAGGAGGTGCAGGTTCCAGAGGT 360
 DB 301 GGAATTGTAAATGAAATTTGGTGGAGCGCTTTTAGTCAGGAGGTGCAGGTTCCAGAGGT 360
 QY 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCAGAAATGTTCACTCAGAGATGTACAGTTTG 420
 DB 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCAGAAATGTTCACTCAGAGATGTACAGTTTG 420
 QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAAATTTTATTAATGCAATTGAA 480
 DB 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAAATTTTATTAATGCAATTGAA 480
 QY 481 ACCATGCCCTATGTTAAGAAAAAGAGATTTGGCCCTTGCATGATGATGAGATGAGAAAA 540
 DB 481 ACCATGCCCTATGTTAAGAAAAAGAGATTTGGCCCTTGCATGATGATGAGATGAGAAAA 540
 QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTTCTCTCTGTAGAGAGGTTTCTTCTCAGGA 600
 DB 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTTCTCTCTGTAGAGAGGTTTCTTCTCAGGA 600
 QY 601 TCTTTTGTCTGCTATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGAGTCTACTTTTTC 660
 DB 601 TCTTTTGTCTGCTATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGAGTCTACTTTTTC 660

```
QY 661 AATGAACATCATCAGCAGAGATGAAGACATTTCACTGTGACTTTTGCCTTGCCTGATGTTCCAA 720
Db 661 AATGAACATCATCAGCAGAGATGAAGACATTTCACTGTGACTTTTGCCTTGCCTGATGTTCCAA 720
QY 721 TACTTAGTAAATAAGCTTTTCAAGAAAGGTCAGGAGATCAATGTTGATGCTGTCAAA 780
Db 721 TACTTAGTAAATAAGCTTTTCAAGAAAGGTCAGGAGATCAATGTTGATGCTGTCAAA 780
QY 781 ATTGAGCAGGAGTTTAAACAGAGCCCTTGCAGTTGGCTCATTTGGAATGAATTCATTT 840
Db 781 ATTGAGCAGGAGTTTAAACAGAGCCCTTGCAGTTGGCTCATTTGGAATGAATTCATTT 840
QY 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTGTGGAACCTTGGATTTCTCA 900
Db 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTGTGGAACCTTGGATTTCTCA 900
QY 901 AAGGTTTTTTCAGGCGAGAAATCCTTTTGATTTTATGAAACATTTCTTTAGAGGAAAA 960
Db 901 AAGGTTTTTTCAGGCGAGAAATCCTTTTGATTTTATGAAACATTTCTTTAGAGGAAAA 960
QY 961 ACAATTTCTTTCAGAAACAGTTTTCAGAGTATCAGCGTTTTCAGAGTATGCGAGAAACC 1020
Db 961 ACAATTTCTTTCAGAAACAGTTTTCAGAGTATCAGCGTTTTCAGAGTATGCGAGAAACC 1020
QY 1021 ACAGATAACGTCCTTACACCTTTGGATGCGAGATTTT 1053
Db 1021 ACAGATAACGTCCTTACACCTTTGGATGCGAGATTTT 1053

RESULT 2
US-10-698-228-4
; Sequence 4, Application US/10698228
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; PRIOR FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 4
; LENGTH: 1081
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-4

Query Match 100.0%; Score 1053; DB 58; Length 1081;
Best Local Similarity 100.0%; Pred. No. 6.2e-276;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGGGACCCGGAAGGCGGAGCGGCTGGATCAGGATCAGAGATCATTTCTCA 60
Db 20 ATGGGGGACCCGGAAGGCGGAGCGGCTGGATCAGGATCAGAGATCATTTCTCA 79
QY 61 GACACCAAGAAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 80 GACACCAAGAAAGTGAATAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 139
QY 121 CGGTTTGTCACTTTCCAAATCCAGTACCCCTGATATTTGGAAATGTTAAACAGGCACAG 180
Db 140 CGGTTTGTCACTTTCCAAATCCAGTACCCCTGATATTTGGAAATGTTAAACAGGCACAG 199
QY 181 GCTTCCTCTGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAAACAAG 240
```

```
Db 200 GCTTCCTCTGAGCAGCAGAAAGAGTTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 259
QY 241 CTTTAAAGCAGATCAGAGTACTTCTCATCTCTCACATCTTAGCCTTTTTCAGCAGCAGTCA 300
Db 260 CTTTAAAGCAGATCAGAGTACTTCTCATCTCTCACATCTTAGCCTTTTTCAGCAGCAGTCA 319
QY 301 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTCAGGTTCCAGAGGCT 360
Db 320 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTCAGGTTCCAGAGGCT 379
QY 361 CGCTGTTTCTATCGCTTTTCAAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTCAGTTTG 420
Db 380 CGCTGTTTCTATCGCTTTTCAAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTCAGTTTG 439
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAAATGCAATGAA 480
Db 440 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAAATGCAATGAA 499
QY 481 ACCATGCCCTATCTTAAGAAAGAGAGATTTGGSCCTTTCGATCGATAGCAGATAGAAA 540
Db 500 ACCATGCCCTATCTTAAGAAAGAGAGATTTGGSCCTTTCGATCGATAGCAGATAGAAA 559
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCCCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 600
Db 560 TCTACTTTTGGGAAAGAGTGGTGGCCCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 619
QY 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGGTCCTTATGCCAGAGTCACTTTTTC 660
Db 620 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGGTCCTTATGCCAGAGTCACTTTTTC 679
QY 661 AATGAACATCATCAGCAGAGATGAAGACATTTCACTGTGACTTTTGCCTTGCCTGATGTTCCAA 720
Db 680 AATGAACATCATCAGCAGAGATGAAGACATTTCACTGTGACTTTTGCCTTGCCTGATGTTCCAA 739
QY 721 TACTTAGTAAATAAGCTTTTCAAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 780
Db 740 TACTTAGTAAATAAGCTTTTCAAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 799
QY 781 ATTGAGCAGGAGTTTAAACAGAGCCCTTGCAGTTTGCCTCATTTGGAATGAATTCATTC 840
Db 800 ATTGAGCAGGAGTTTAAACAGAGCCCTTGCAGTTTGCCTCATTTGGAATGAATTCATTC 859
QY 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTGTGGAACCTTGGATTTCTCA 900
Db 860 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTGTGGAACCTTGGATTTCTCA 919
QY 901 AAGGTTTTTTCAGGCGAGAAATCCTTTTGATTTTATGAAACATTTCTTTAGAGGAAAA 960
Db 920 AAGGTTTTTTCAGGCGAGAAATCCTTTTGATTTTATGAAACATTTCTTTAGAGGAAAA 979
QY 961 ACAATTTCTTTCAGAAACAGTTTTCAGAGTATCAGCGTTTTCAGAGTATGCGAGAAACC 1020
Db 980 ACAATTTCTTTCAGAAACAGTTTTCAGAGTATCAGCGTTTTCAGAGTATGCGAGAAACC 1039
QY 1021 ACAGATAACGTCCTTACACCTTTGGATGCGAGATTTT 1053
Db 1040 ACAGATAACGTCCTTACACCTTTGGATGCGAGATTTT 1072
```

```
RESULT 3
US-10-698-228-2
; Sequence 2, Application US/10698228
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; PRIOR FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
```

;; PRIOR APPLICATION NUMBER: JP 11-192391
;; PRIOR FILING DATE: 1999-07-06
;; PRIOR APPLICATION NUMBER: JP 2000-017770
;; PRIOR FILING DATE: 2000-01-21
;; NUMBER OF SEQ ID NOS: 14
;; SEQ ID NO 2
;; LENGTH: 1053
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION:
US-10-698-228-2

Query Match 99.8%; Score 1051.4; DB 58; Length 1053;
Best Local Similarity 99.9%; Pred. No. 1.7e-275;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGACCCGGAAGCGCGGCGGCTGGATCAGAGATGAGATCATCTTCA 60
DB 1 ATGGCGACCCGGAAGCGCGGCGGCTGGATCAGAGATGAGATCATCTTCA 60
QY 61 GACACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGATTCTCG 120
DB 61 GACACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGATTCTCG 120
QY 121 CGTTTGTGTCATCTTCCCAATCCAGTACCCTGATATTTGAAAATGATATAACAGGCACAG 180
DB 121 CGTTTGTGTCATCTTCCCAATCCAGTACCCTGATATTTGAAAATGATATAACAGGCACAG 180
QY 181 GCTTCCTTCTGACAGCAGAGAGGTGACTTATCAAGAGATCTCCCTCACTGGAAACAAG 240
DB 181 GCTTCCTTCTGACAGCAGAGAGGTGACTTATCAAGAGATCTCCCTCACTGGAAACAAG 240
QY 241 CTTAAGCAGATGAGAAGTACTTCTCTCATCTTCACTTACCTTTTTCAGCCAGTGTAT 300
DB 241 CTTAAGCAGATGAGAAGTACTTCTCTCATCTTCACTTACCTTTTTCAGCCAGTGTAT 300
QY 301 GGAATTTGAATGAATAATTTGGTGAGCGCTTTAGTCAGAGGTGCGAGGTCCAGAGGCT 360
DB 301 GGAATTTGAATGAATAATTTGGTGAGCGCTTTAGTCAGAGGTGCGAGGTCCAGAGGCT 360
QY 361 CGCTGTTTCTATGGCTTTCAAAATTCATCGAAGATTTTCACTCAGAGATGTACAGTTTG 420
DB 361 CGCTGTTTCTATGGCTTTCAAAATTCATCGAAGATTTTCACTCAGAGATGTACAGTTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA 480
DB 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA 480
QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATTCGGCTTCGGATGGATAGCAGATAGAAA 540
DB 481 ACCATGCCCTATGTTAAGAAAAGCAGATTCGGCTTCGGATGGATAGCAGATAGAAA 540
QY 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTGTAAGAGAGATTCTTCTCAGGA 600
DB 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTGTAAGAGAGATTCTTCTCAGGA 600
QY 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAAAGAGAGTCTTATGCCAGGACTCATTTTCC 660
DB 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAAAGAGAGTCTTATGCCAGGACTCATTTTCC 660
QY 661 AATGAACATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
DB 661 AATGAACATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
QY 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGTGAGGAGATCATTTGCTGCTGTCAAA 780
DB 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGTGAGGAGATCATTTGCTGCTGTCAAA 780
QY 781 ATTGACGAGGAGTTTTTAA CAGAAGCCTTGGCAGTTGGCCTCATTTGGAATGAATTGCAAT 840
DB 781 ATTGACGAGGAGTTTTTAA CAGAAGCCTTGGCAGTTGGCCTCATTTGGAATGAATTGCAAT 840

QY 841 TTGATGAAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
DB 841 TTGATGAAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
QY 901 AAGGTTTTTCAGGCAGAAAATCCTTTTGTGATTTTATGAAAACATTCTTTAGAAAGGAAA 960
DB 901 AAGGTTTTTCAGGCAGAAAATCCTTTTGTGATTTTATGAAAACATTCTTTAGAAAGGAAA 960
QY 961 ACAAAATTTCTTTGAGAAAACAGTTTTCAGAGTATCAGCGTTTGTGAGTTATGCGCAGAAACC 1020
DB 961 ACAAAATTTCTTTGAGAAAACAGTTTTCAGAGTATCAGCGTTTGTGAGTTATGCGCAGAAACC 1020
QY 1021 ACAGATAACGTCCTTTCACCTTTGGATGAGATTTT 1053
DB 1021 ACAGATAACGTCCTTTCACCTTTGGATGAGATTTT 1053

RESULT 4
US-09-787-491B-28
; Sequence 28, Application US/09787491B
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; TANG, Y. Tom;
; APPLICANT: CORLEY, Neil C.; GUEGLER, Karl J.;
; APPLICANT: GORGONE, Gina A.; PATTERSON, Chandra;
; APPLICANT: HILLMAN, Jennifer L.; BAUGHN, Mariah R.;
; APPLICANT: LAL, Preeti; AZIMZAI, Valda;
; APPLICANT: YUE, Henry; YANG, Junning
; TITLE OF INVENTION: RNA-ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0600 USN
; CURRENT APPLICATION NUMBER: US/09/787,491B
; CURRENT FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: PCT/US99/21688
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: US 60/128,660
; PRIOR FILING DATE: 1999-04-08
; PRIOR APPLICATION NUMBER: US 60/069,391
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: US 60/183,025
; PRIOR FILING DATE: 1998-09-22
; PRIOR APPLICATION NUMBER: US 60/155,246
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: US 09/158,720
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PERL Program
; SEQ ID NO 28
; LENGTH: 2596
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 2604449CBI
US-09-787-491B-28

Query Match 99.8%; Score 1051.4; DB 34; Length 2596;
Best Local Similarity 99.9%; Pred. No. 2.4e-275;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAGCGCGGCGGCTGGATCAGAGATGAGATCATCTTCA 60
DB 56 ATGGGCGACCCGGAAGCGCGGCGGCTGGATCAGAGATGAGATCATCTTCA 115
QY 61 GACACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGATTCTCGC 120
DB 116 GACACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGATTCTCGC 175
QY 121 CGTTTGTGTCATCTTCCCAATCCAGTACCCTGATATTTGAAAATGATATAACAGGCACAG 180
DB 176 CGTTTGTGTCATCTTCCCAATCCAGTACCCTGATATTTGAAAATGATATAACAGGCACAG 235
QY 181 GCTTCCTTCTGACAGCAGAGAGGTGACTTATCAAGAGATCTCCCTCACTGGAAACAAG 240
DB 236 GCTTCCTTCTGACAGCAGAGAGGTGACTTATCAAGAGATCTCCCTCACTGGAAACAAG 295

```
QY 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTCTTACCTTTTTCAGCCAGTGAT 300
Db 296 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTCTTACCTTTTTCAGCCAGTGAT 355
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGAGGCT 360
Db 356 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGAGGCT 415
QY 361 CGCTGTTTCTATGGCTTCCAAATTCATCGAGAAATGTTCACTCAGAGTGTACGTTTG 420
Db 416 CGCTGTTTCTATGGCTTCCAAATTCATCGAGAAATGTTCACTCAGAGTGTACGTTTG 475
QY 421 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAA 480
Db 476 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAA 535
QY 481 ACCATGCCCTATGTTAAGAAAGAGAGATGAGGCTTGGCTTGGATGAGATAGAA 540
Db 536 ACCATGCCCTATGTTAAGAAAGAGAGATGAGGCTTGGCTTGGATGAGATAGAA 595
QY 541 TCTACTTTTGGGAAGAGTGGTGGCTTGGCTTGGCTGAGAGAGTGTCTTCTCAGGA 600
Db 596 TCTACTTTTGGGAAGAGTGGTGGCTTGGCTTGGCTGAGAGAGTGTCTTCTCAGGA 655
QY 601 TCTTTTGGCTGCTATATTTCTGGCTTAAAGAAAGAGAGTCTTATGCGAGAGTCTTCTTCC 660
Db 656 TCTTTTGGCTGCTATATTTCTGGCTTAAAGAAAGAGAGTCTTATGCGAGAGTCTTCTTCC 715
QY 661 AATGAATCTCATCAGCAGATGAAGGACTTCACTGTGACTTTTGGCTTGGCTGATGTTCCAA 720
Db 716 AATGAATCTCATCAGCAGATGAAGGACTTCACTGTGACTTTTGGCTTGGCTGATGTTCCAA 775
QY 721 TACTTAGTAAATAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTAAGTGTCTCA 780
Db 776 TACTTAGTAAATAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTAAGTGTCTCA 835
QY 781 ATTGAGCAGAGTCTTAAACAGAGCCTTGGCAGTGTGGCTTCACTGGAATGGAATTCGATT 840
Db 836 ATTGAGCAGAGTCTTAAACAGAGCCTTGGCAGTGTGGCTTCACTGGAATGGAATTCGATT 895
QY 841 TTGATGAACAGTACATTTAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 900
Db 896 TTGATGAACAGTACATTTAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 955
QY 901 AAGGTTTTTTCAGCAGAAATCTTTTGAATTTTATGGAAGAAATTTCTTTAGAGGAA 960
Db 956 AAGGTTTTTTCAGCAGAAATCTTTTGAATTTTATGGAAGAAATTTCTTTAGAGGAA 1015
QY 961 ACAATTTCTTTGAGAAAGAGTTCAGAGTATCAGCGTTTTCAGGTTATGCGAGAAACC 1020
Db 1016 ACAATTTCTTTGAGAAAGAGTTCAGAGTATCAGCGTTTTCAGGTTATGCGAGAAACC 1075
QY 1021 ACAGATAAGCTCTTCACTTGGATGAGATTTT 1053
Db 1076 ACAGATAAGCTCTTCACTTGGATGAGATTTT 1108
```

RESULT 5

```
US-60-128-660-20
; Sequence 20, Application US/60128660
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Tang, Y. Tom
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Yang, Junming
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: RNA-ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0682 P
; CURRENT APPLICATION NUMBER: US/60128,660
; CURRENT FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 26
```

```
; SOFTWARE: PERL Program
; SEQ ID NO 20
; LENGTH: 2596
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 2604449
US-60-128-660-20
```

```
Query Match 99.8%; Score 1051.4; DB 81; Length 2596;
Best Local Similarity 99.9%; Pred. No. 2.4e-275;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 56 ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 115
QY 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 116 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 175
QY 121 CGGTTTGTCTATCTTCCAAATCCAGTACCTGTATATTTGGAAATGTATAAACAGGCAAG 180
Db 176 CGGTTTGTCTATCTTCCAAATCCAGTACCTGTATATTTGGAAATGTATAAACAGGCAAG 235
QY 181 GCTTCTCTTGGACAGCAGAGAGGTTGACTTATCAAGAGGATCTCCCTCAGTGGAAACAG 240
Db 236 GCTTCTCTTGGACAGCAGAGAGGTTGACTTATCAAGAGGATCTCCCTCAGTGGAAACAG 295
QY 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTCTTAGCCTTTTTCAGCCAGTGAT 300
Db 296 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTCTTAGCCTTTTTCAGCCAGTGAT 355
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGAGGCT 360
Db 356 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGAGGCT 415
QY 361 CGCTGTTTCTATGGCTTCCAAATTCATCGAGAAATGTTCACTCAGAGTGTACGTTTG 420
Db 416 CGCTGTTTCTATGGCTTCCAAATTCATCGAGAAATGTTTATTAATGCAATTTGAA 475
QY 421 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAA 480
Db 476 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAA 535
QY 481 ACCATGCCCTATGTTAAGAAAGAGAGTGTGGCCTTGGCTTGGATGAGATAGAA 540
Db 536 ACCATGCCCTATGTTAAGAAAGAGAGTGTGGCCTTGGCTTGGATGAGATAGAA 595
QY 541 TCTACTTTTGGGAAGAGTGGTGGCTTGGCTTGGCTGAGAGGTTTCTTCTCAGGA 600
Db 596 TCTACTTTTGGGAAGAGTGGTGGCTTGGCTTGGCTGAGAGGTTTCTTCTCAGGA 655
QY 601 TCTTTTGGCTGCTATATTTCTGGCTTAAAGAAAGAGAGTCTTATGCGAGAGTCTTATGCGAGAGTCTTCTTCC 660
Db 656 TCTTTTGGCTGCTATATTTCTGGCTTAAAGAAAGAGAGTCTTATGCGAGAGTCTTATGCGAGAGTCTTCTTCC 715
QY 661 AATGAATCTCATCAGCAGATGAAGGACTTCACTGTGACTTTTGGCTTGGCTGATGTTCCAA 720
Db 716 AATGAATCTCATCAGCAGATGAAGGACTTCACTGTGACTTTTGGCTTGGCTGATGTTCCAA 775
QY 721 TACTTAGTAAATAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTAAGTGTCTCA 780
Db 776 TACTTAGTAAATAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTAAGTGTCTCA 835
QY 781 ATTGAGCAGAGTCTTAAACAGAGCCTTGGCAGTGTGGCTTCACTGGAATGGAATTCGATT 840
Db 836 ATTGAGCAGAGTCTTAAACAGAGCCTTGGCAGTGTGGCTTCACTGGAATGGAATTCGATT 895
QY 841 TTGATGAACAGTACATTTAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 900
Db 896 TTGATGAACAGTACATTTAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 955
```

```
QY 901 AAGTGTTCAGCAGAGAAATCCTTTGATTTATGAAACATTTCTTTAGAGGAAA 960
Db |||||||
QY 956 AAGTGTTCAGCAGAGAAATCCTTTGATTTATGAAACATTTCTTTAGAGGAAA 1015
Db |||||||
QY 961 ACAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC 1020
Db |||||||
QY 1016 ACAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC 1075
Db |||||||
QY 1021 ACAGATAAGCTCTTACCTTTGGATGAGATTTT 1053
Db |||||||
QY 1076 ACAGATAAGCTCTTACCTTTGGATGAGATTTT 1108
Db |||||||

RESULT 6
US-10-170-235-20856
; Sequence 20856, Application US/10170235
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN
; TITLE OF INVENTION: TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF
; FILE REFERENCE: CL001380
; CURRENT APPLICATION NUMBER: US/10/170,235
; CURRENT FILING DATE: 2003-03-17
; NUMBER OF SEQ ID NOS: 42514
; SEQ ID NO 20856
; LENGTH: 3397
; TYPE: DNA
; ORGANISM: HUMAN
US-10-170-235-20856

Query Match 99.8%; Score 1051.4; DB 49; Length 3397;
Best Local Similarity 99.9%; Pred. No. 2.7e-275;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATGAGAGATCATCTTCA 60
Db |||||||
QY 501 ATGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATGAGAGATCATCTTCA 560
Db |||||||
QY 61 GACACCAAGCAAGTGAATAAGTCAATGAAGCCACTCTTAAGAAAGTTCCTCG 120
Db |||||||
QY 561 GACACCAAGCAAGTGAATAAGTCAATGAAGCCACTCTTAAGAAAGTTCCTCG 620
Db |||||||
QY 121 CGGTTTGTATCTTTCCTCAATCCAGTACCTGATATTGGAAATGATATAAGGACAG 180
Db |||||||
QY 621 CGGTTTGTATCTTTCCTCAATCCAGTACCTGATATTGGAAATGATATAAGGACAG 680
Db |||||||
QY 181 GCTTCCTCTTGACAGCAGAGAGGTGACTTATCAAGGATCTCCCTCACTGGAACAAG 240
Db |||||||
QY 681 GCTTCCTCTTGACAGCAGAGAGGTGACTTATCAAGGATCTCCCTCACTGGAACAAG 740
Db |||||||
QY 241 CTTAAGCGAGATGAGAGTACTTCTCTCAGATCTTAGCCCTTTTTCAGCCAGTGTAT 300
Db |||||||
QY 741 CTTAAGCGAGATGAGAGTACTTCTCTCAGATCTTAGCCCTTTTTCAGCCAGTGTAT 800
Db |||||||
QY 301 GGAATTTGTAATGAAATTTTGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
Db |||||||
QY 801 GGAATTTGTAATGAAATTTTGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 860
Db |||||||
QY 361 CGCTGTTCTATGCGCTTCAAAATTCATCGAGATGTTCACTCAGAGATGTACAGTTTG 420
Db |||||||
QY 861 CGCTGTTCTATGCGCTTCAAAATTCATCGAGATGTTCACTCAGAGATGTACAGTTTG 920
Db |||||||
QY 421 CTGATAGACATTTATCATGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 480
Db |||||||
QY 921 CTGATAGACATTTATCATGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 980
Db |||||||
QY 481 ACCATGCCCTATTTAAGAAAGAGAGATGGCCCTTGGATGAGATGAGATGAGATGAAA 540
Db |||||||
QY 981 ACCATGCCCTATTTAAGAAAGAGAGATGGCCCTTGGATGAGATGAGATGAGATGAAA 1040
Db |||||||
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTTGTCTGTAGAGAGAGTTCCTCTCAGGA 600
Db |||||||
QY 1041 TCTACTTTTGGGGAAGAGTGGTGGCTTTTGTCTGTAGAGAGAGTTCCTCTCAGGA 1100
Db |||||||
```

```
QY 601 TCTTTTGTGCTATATTCTGGCTAAAGAGAGAGGTCTTATGCCAGACTCACATTTTCC 660
Db |||||||
QY 1101 TCTTTTGTGCTATATTCTGGCTAAAGAGAGAGGTCTTATGCCAGACTCACATTTTCC 1160
Db |||||||
QY 661 AATGAACTCATCAGCAGAGATGAAGACTTTCACGTGACATTTGCTTCCCTGATGTTCAA 720
Db |||||||
QY 1161 AATGAACTCATCAGCAGAGATGAAGACTTTCACGTGACATTTGCTTCCCTGATGTTCAA 1220
Db |||||||
QY 721 TACTTAGTAAATAGCCCTTCAGAAAGAGGTTCAGGAGATCATTTGATGCTGTCAA 780
Db |||||||
QY 1221 TACTTAGTAAATAGCCCTTCAGAAAGAGGTTCAGGAGATCATTTGATGCTGTCAA 1280
Db |||||||
QY 781 ATTGACGAGAGTTCCTTAAACAGAAAGCTTCGCCAGTTGGCCTCATTTGCAATGCAATT 840
Db |||||||
QY 1281 ATTGACGAGAGTTCCTTAAACAGAAAGCTTCGCCAGTTGGCCTCATTTGCAATGCAATT 1340
Db |||||||
QY 841 TTGATGAAACAGTACATGAGTTTGTAGCTGACAGATTTACTTTGGAACTTTGGATTTCA 900
Db |||||||
QY 1341 TTGATGAAACAGTACATGAGTTTGTAGCTGACAGATTTACTTTGGAACTTTGGATTTCA 1400
Db |||||||
QY 901 AAGTGTTCAGCAGAGAAATCCTTTGATTTATGAAACATTTCTTTAGAGGAAA 960
Db |||||||
QY 1401 AAGTGTTCAGCAGAGAAATCCTTTGATTTATGAAACATTTCTTTAGAGGAAA 1460
Db |||||||
QY 961 ACAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC 1020
Db |||||||
QY 1461 ACAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC 1520
Db |||||||
QY 1021 ACAGATAAGCTCTTACCTTTGGATGAGATTTT 1053
Db |||||||
QY 1521 ACAGATAAGCTCTTACCTTTGGATGAGATTTT 1553
Db |||||||
```

RESULT 7

```
PCT-US02-18947-71
; Sequence 71, Application PC/TUS0218947
; GENERAL INFORMATION:
; APPLICANT: Rosetta Inpharmatics
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-228
; CURRENT APPLICATION NUMBER: PCT/US02/18947
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 71
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AB036063
; DATABASE ENTRY DATE: 2001-06-18
PCT-US02-18947-71
```

Query Match 99.8%; Score 1051.4; DB 1; Length 4955;
Best Local Similarity 99.9%; Pred. No. 3.2e-275;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 ATGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGAGATGAGATCATCTTCA 60
Db |||||||
QY 245 ATGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGAGATGAGATCATCTTCA 304
Db |||||||
QY 61 GACACCAAGCAAGTGAATAAGTCAATGAAGCCACTCTTAAGAAAGTTCCTCGC 120
Db |||||||
QY 305 GACACCAAGCAAGTGAATAAGTCAATGAAGCCACTCTTAAGAAAGTTCCTCGC 364
Db |||||||
QY 121 CGGTTTGTATCTTTCCTCAATCCAGTACCTGATATTGGAAATGATATAAGGACAG 180
Db |||||||
QY 365 CGGTTTGTATCTTTCCTCAATCCAGTACCTGATATTGGAAATGATATAAGGACAG 424
Db |||||||
QY 181 GCTTCCTCTTGAGAGAGAGTGGTGGCTTTTGTCTGTAGAGAGTTCCTCTCAGGA 240
Db |||||||
```

```
Db 425 GCTTCCTCTGACAGCAGAGAGGTCGACTTATCAAGAGATCTCCCTCACTGGAACAAG 484
Qy 241 CTTAAGCAGATGAGAGTACTTCACTCTCATCTCTAGCCCTTTTTCGAGCCAGTGAT 300
Db 485 CTTAAGCAGATGAGAGTACTTCACTCTCATCTCTAGCCCTTTTTCGAGCCAGTGAT 544
Qy 301 GGAATTTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
Db 545 GGAATTTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 604
Qy 361 CGCTGTTCTATGCTTTCAAAATTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG 420
Db 605 CGCTGTTCTATGCTTTCAAAATTCATCGAAGATGTTCACTCAGAGATGTACAGTTTG 664
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAATGCAATTTGAA 480
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAATGCAATTTGAA 724
Qy 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGGATGGATAGCAGATAGAAA 540
Db 725 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGGATGGATAGCAGATAGAAA 784
Qy 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGTGTAGAGAGTGTCTTCTCAGGA 600
Db 785 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGTGTAGAGAGTGTCTTCTCAGGA 844
Qy 601 TCTTTTGTCTATATTTCTGGCTTAAGAAAGAGAGTCTTATGCGCAGGACTCACTTTTTC 660
Db 845 TCTTTTGTCTATATTTCTGGCTTAAGAAAGAGAGTCTTATGCGCAGGACTCACTTTTTC 904
Qy 661 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGGCTGTATGTTCCAA 720
Db 905 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGGCTGTATGTTCCAA 964
Qy 721 TACTTAGTAAATAAGCCTTTCAGAGAAAGGCTCAGGAGATCATTTGTCATGCTGTCAAA 780
Db 965 TACTTAGTAAATAAGCCTTTCAGAGAAAGGCTCAGGAGATCATTTGTCATGCTGTCAAA 1024
Qy 781 ATTGAGCAGGAGTTTAAACAGAGCCTTTCAGAGCCTTTCAGAGTGGCTCATTTGGAATGCAATT 840
Db 1025 ATTGAGCAGGAGTTTAAACAGAGCCTTTCAGAGCCTTTCAGAGTGGCTCATTTGGAATGCAATT 1084
Qy 841 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTAATTTGTTGAACTTGGATTTCTCA 900
Db 1085 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTAATTTGTTGAACTTGGATTTCTCA 1144
Qy 901 AAGGTTTTTCAGGCAAGAAAAATCCTTTGATTTTATGGAACAATTTCTTTAGAGGAAAA 960
Db 1145 AAGGTTTTTCAGGCAAGAAAAATCCTTTGATTTTATGGAACAATTTCTTTAGAGGAAAA 1204
Qy 961 ACAAATTTCTTTGAGAAACAGGTTTCAGAGTATCAGCGTTTTCAGAGTTATGCGAGAAACC 1020
Db 1205 ACAAATTTCTTTGAGAAACAGGTTTCAGAGTATCAGCGTTTTCAGAGTTATGCGAGAAACC 1264
Qy 1021 ACAGATAACGCTTTCACCTTGGATGCAAGATTTT 1053
Db 1265 ACAGATAACGCTTTCACCTTGGATGCAAGATTTT 1297
```

RESULT 8

```
; Sequence 71, Application PC/TUS0401100
; GENERAL INFORMATION:
; APPLICANT: Rosetta Inpharmatics LLC
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; CURRENT APPLICATION NUMBER: PCT/US04/01100
; PRIOR FILING DATE: 2004-01-15
; PRIOR APPLICATION NUMBER: 10/342,887
; PRIOR FILING DATE: 2003-01-15
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 71
; LENGTH: 4955
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US04-01100-71
```

```
Query Match 99.8%; Score 1051.4; DB 2; Length 4955;
Best Local Similarity 99.9%; Pred. No. 3.2e-275;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 ATGGCGCACCGGAAAGCGCCGGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 245 ATGGCGCACCGGAAAGCGCCGGGCTGGATCAGGATGAGAGATCATCTTCA 304
Qy 61 GACACCAACCAAGCTGAAATAAAGTCAAAATCAAGAGCCACTCTTAAGAAAGAGTTCTCCG 120
Db 305 GACACCAACCAAGCTGAAATAAAGTCAAAATCAAGAGCCACTCTTAAGAAAGAGTTCTCCG 364
Qy 121 CGGTTTGTCTATCTTTCCAAATCCAGTACCTCATATTTGGAAAAATGTATAAACAGGCACAG 180
Db 365 CGGTTTGTCTATCTTTCCAAATCCAGTACCTCATATTTGGAAAAATGTATAAACAGGCACAG 424
Qy 181 GCTTCCTTCTGGACAGCAGAGAGGTGACATTATCAAAAGGATCTCCCTCATCGGAACAAG 240
Db 425 GCTTCCTTCTGGACAGCAGAGAGGTGACATTATCAAAAGGATCTCCCTCATCGGAACAAG 484
Qy 241 CTTAAAGCAGATGAGAGTACTTCTCTCATCTCTACATCTTTAGCCTTTTTCAGCCAGTGAT 300
Db 485 CTTAAAGCAGATGAGAGTACTTCTCTCATCTCTACATCTTTAGCCTTTTTCAGCCAGTGAT 544
Qy 301 GGAATTTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
Db 545 GGAATTTGTAATGAAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 604
Qy 361 CGCTGTTTCTATGCTTTCAAAATTCATCAGAGATGTTCACCTCAGAGATGTACAGTTTG 420
Db 605 CGCTGTTTCTATGCTTTCAAAATTCATCAGAGATGTTCACCTCAGAGATGTACAGTTTG 664
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAATGCAATTTGAA 480
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGGAATTTTATTTAATGCAATTTGAA 724
Qy 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGGATGGATAGCAGATAGAAA 540
Db 725 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGGATGGATAGCAGATAGAAA 784
Qy 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAGAGGAGTTTCTTCTCAGGA 600
Db 785 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAGAGGAGTTTCTTCTCAGGA 844
Qy 601 TCTTTTGTCTGTATATTTCTGGCTTAAGAAAGAGGCTTATGCGCAGGACTCACTTTTTC 660
Db 845 TCTTTTGTCTGTATATTTCTGGCTTAAGAAAGAGGCTTATGCGCAGGACTCACTTTTTC 904
Qy 661 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGGCTGTATGTTCCAA 720
Db 905 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGGCTGTATGTTCCAA 964
Qy 721 TACTTAGTAAATAAGCCTTTCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAAA 780
Db 965 TACTTAGTAAATAAGCCTTTCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAAA 1024
Qy 781 ATTGAGCAGGAGTTTAAACAGAGCCTTTCAGAGCCTTTCAGAGTGGCTCATTTGGAATGCAATT 840
Db 1025 ATTGAGCAGGAGTTTAAACAGAGCCTTTCAGAGCCTTTCAGAGTGGCTCATTTGGAATGCAATT 1084
Qy 841 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTAATTTGTTGAACTTGGATTTCTCA 900
Db 1085 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTAATTTGTTGAACTTGGATTTCTCA 1144
Qy 901 AAGGTTTTTCAGGCAAGAAAAATCCTTTGATTTTATGGAACAATTTCTTTAGAGGAAAA 960
Db 1145 AAGGTTTTTCAGGCAAGAAAAATCCTTTGATTTTATGGAACAATTTCTTTAGAGGAAAA 1204
Qy 961 ACAAATTTCTTTGAGAAACAGGTTTCAGAGTATCAGCGTTTTCAGAGTTATGCGAGAAACC 1020
Db 1205 ACAAATTTCTTTGAGAAACAGGTTTCAGAGTATCAGCGTTTTCAGAGTTATGCGAGAAACC 1264
Qy 1021 ACAGATAACGCTTTCACCTTGGATGCAAGATTTT 1053
Db 1265 ACAGATAACGCTTTCACCTTGGATGCAAGATTTT 1297
```


QY 661 AATGAACATCATCAGCAGAGATGAAGACATTTCACTGTGACTTTGCTTGCTGCTGATGTTCCAA 720
DB 905 AATGAACATCATCAGCAGAGATGAAGACATTTCACTGTGACTTTGCTTGCTGCTGATGTTCCAA 964
QY 721 TACTTAGTAATAAGCTTTTCAAGAAAGGGTCAGGAGATCATTTGATGCTGTCAAA 780
DB 965 TACTTAGTAATAAGCTTTTCAAGAAAGGGTCAGGAGATCATTTGATGCTGTCAAA 1024
QY 781 ATTGACGAGAGTTTAAACAGAGCTTGCAGTTGGCTCATTTGGAATGAATTCATTT 840
DB 1025 ATTGACGAGAGTTTAAACAGAGCTTGCAGTTGGCTCATTTGGAATGAATTCATTT 1084
QY 841 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGAACTTTGATTTCTCA 900
DB 1085 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGAACTTTGATTTCTCA 1144
QY 901 AAGGTTTTTTCAGGAGAGAAATCTTTTGATTTTATGGAAGAAATCTTTTGAAGGAAAA 960
DB 1145 AAGGTTTTTTCAGGAGAGAAATCTTTTGATTTTATGGAAGAAATCTTTTGAAGGAAAA 1204
QY 961 ACAAAATTTCTTGAAGAAAGCTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC 1020
DB 1205 ACAAAATTTCTTGAAGAAAGCTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC 1264
QY 1021 ACAGATAACGCTCTTCACTTGGATGAGATTTT 1053
DB 1265 ACAGATAACGCTCTTCACTTGGATGAGATTTT 1297

RESULT 12

US-10-990-328-49
; Sequence 49, Application US/10990328
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 1849
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-990-328-49

Query Match 99.8%; Score 1051; DB 67; Length 1849;
Best Local Similarity 99.8%; Pred. No. 2.7e-275;
Matches 1051; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAGCGCGGAGCGGCTGGATCAGATCAGAGATCATCTTCA 60
DB 85 ATGGGCGACCCGGAAGCGCGGAGCGGCTGGATCAGATCAGAGATCATCTTCA 144
QY 61 GACACCAACGAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
DB 145 GACACCAACGAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 204
QY 121 CGTTTGTGATCTTCCATCCAGTACCTGATATTTGGAATGATGTAACAGGACAG 180
DB 205 CGTTTGTGATCTTCCATCCAGTACCTGATATTTGGAATGATGTAACAGGACAG 264
QY 181 GCTTCTCTCTGACAGCAGAGAGTTGACTTATCAAGATCTCCCTCACTGGAACAAG 240
DB 265 GCTTCTCTCTGACAGCAGAGAGTTGACTTATCAAGATCTCCCTCACTGGAACAAG 324
QY 241 CTTAAGCAGATGAGAAGTACTTTCATCTCTCATCTTTCAGCCCTTTTTCAGCCAGTGAT 300
DB 325 CTTAAGCAGATGAGAAGTACTTTCATCTCTCATCTTTCAGCCCTTTTTCAGCCAGTGAT 384

QY 301 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
DB 385 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 444
QY 361 CGCTGTTTCTATGGCTTTCAAAATTCATCGAAGATTTCACTCAGAGATGTAAGTTTG 420
DB 445 CGCTGTTTCTATGGCTTTCAAAATTCATCGAAGATTTCACTCAGAGATGTAAGTTTG 504
QY 421 CTGATAGACATTTACATCAGATCCCAAGAAAGGGAATTTTATTAATGCAATTCAA 480
DB 505 CTGATAGACATTTACATCAGATCCCAAGAAAGGGAATTTTATTAATGCAATTCAA 564
QY 481 ACCATGCCCTATCTTAAGAAAAAGCAGATTGGCCCTTGCATGATAGCAGATAGAAA 540
DB 565 ACCATGCCCTATCTTAAGAAAAAGCAGATTGGCCCTTGCATGATAGCAGATAGAAA 624
QY 541 TCTACTTTTGGGGAAGAGTGGCTTTGCTGCTGTAAGAGAGTTTCTTCTCAGGA 600
DB 625 TCTACTTTTGGGGAAGAGTGGCTTTGCTGCTGTAAGAGAGTTTCTTCTCAGGA 684
QY 601 TCTTTTGTCTATATCTGGCTTAAAGAGAGAGGTTCTATGCCAGGACTCACATTTTCC 660
DB 685 TCTTTTGTCTATATCTGGCTTAAAGAGAGAGGTTCTATGCCAGGACTCACATTTTCC 744
QY 661 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCTGATGTTCAA 720
DB 745 AATGAACATCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCTGATGTTCAA 804
QY 721 TACTTAGTAATAAGCTTTTCAAGAAAGGGTCAGGAGATCATTTGATGCTGTCAA 780
DB 805 TACTTAGTAATAAGCTTTTCAAGAAAGGGTCAGGAGATCATTTGATGCTGTCAA 864
QY 781 ATTGACGAGAGTTTAAACAGAGCTTGCAGTTGGCTCATTTGGAATGAATTCATTT 840
DB 865 ATTGACGAGAGTTTAAACAGAGCTTGCAGTTGGCTCATTTGGAATGAATTCATTT 924
QY 841 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGAACTTTGGAATTCCTCA 900
DB 925 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGAACTTTGGAATTCCTCA 984
QY 901 AAGGTTTTTTCAGGAGAGAAATCTTTTGATTTTATGGAAGAAATCTTTTGAAGGAAAA 960
DB 985 AAGGTTTTTTCAGGAGAGAAATCTTTTGATTTTATGGAAGAAATCTTTTGAAGGAAAA 1044
QY 961 ACAAAATTTCTTGAAGAAAGCTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC 1020
DB 1045 ACAAAATTTCTTGAAGAAAGCTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC 1104
QY 1021 ACAGATAACGCTCTTCACTTGGATGAGATTTT 1053
DB 1105 ACAGATAACGCTCTTCACTTGGATGAGATTTT 1137

RESULT 13

US-60-505-218-20
; Sequence 20, Application US/60505218
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CANCER, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001482
; CURRENT APPLICATION NUMBER: US/60/505,218
; CURRENT FILING DATE: 2003-09-24
; NUMBER OF SEQ ID NOS: 22507
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 1849
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-505-218-20

Query Match 99.8%; Score 1051; DB 120; Length 1849;
Best Local Similarity 99.8%; Pred. No. 2.7e-275;

Matches 1051; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATCAGAGATCATCTTCA 60
DB |||||
QY 85 ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATCAGAGATCATCTTCA 144
DB |||||
QY 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
DB |||||
QY 145 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 204
DB |||||
QY 121 CGGTTTGTATCTTCCAAATCCAGTACCTGTATATTTGGAAAATGATATAAAGAGGACAG 180
DB |||||
QY 205 CGGTTTGTATCTTCCAAATCCAGTACCTGTATATTTGGAAAATGATATAAAGAGGACAG 264
DB |||||
QY 181 GCTTCTCTCTGACAGCAGAGAGGTTGACTTATCAAGGATCTCCCTCACTGGACACAG 240
DB |||||
QY 265 GCTTCTCTCTGACAGCAGAGAGGTTGACTTATCAAGGATCTCCCTCACTGGACACAG 324
DB |||||
QY 241 CTTAAAGCAGATCAGAAAGTACTTCTCTCATCTCTAGCCCTTTTTCGAGCCAGTGAT 300
DB |||||
QY 325 CTTAAAGCAGATCAGAAAGTACTTCTCTCATCTCTAGCCCTTTTTCGAGCCAGTGAT 384
DB |||||
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGAGTGAGGTTCCAGAGGCT 360
DB |||||
QY 385 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGAGTGAGGTTCCAGAGGCT 444
DB |||||
QY 361 CGCTGTTCTATGCGCTTTTCAAAATCTCATCGAGATGTTCTCATCAGAGATGACAGTTTG 420
DB |||||
QY 445 CGCTGTTCTATGCGCTTTTCAAAATCTCATCGAGATGTTCTCATCAGAGATGACAGTTTG 504
DB |||||
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATGAA 480
DB |||||
QY 505 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATGAA 564
DB |||||
QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGCCCTTCGCGATGAGATAGAGATAAGAAA 540
DB |||||
QY 565 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGCCCTTCGCGATGAGATAGAGATAAGAAA 624
DB |||||
QY 541 TCTACTTTTGGGAAAAGAGTGTTGGCTTTGCTGCTGATAGAGAGTTTCTTCTCAGGA 600
DB |||||
QY 625 TCTACTTTTGGGAAAAGAGTGTTGGCTTTGCTGCTGATAGAGAGTTTCTTCTCAGGA 684
DB |||||
QY 601 TCTTTTGTGCTATATTCTGCTTAAGAGAGAGGCTTTATGCCAGACTCACTTTTTC 660
DB |||||
QY 685 TCTTTTGTGCTATATTCTGCTTAAGAGAGAGGCTTTATGCCAGACTCACTTTTTC 744
DB |||||
QY 661 AATGAACCTCATCAGCAGAGATGAAGACTTCTACTGTGACTTTGCTTGCCTGATGTTCCAA 720
DB |||||
QY 745 AATGAACCTCATCAGCAGAGATGAAGACTTCTACTGTGACTTTGCTTGCCTGATGTTCCAA 804
DB |||||
QY 721 TACTTAGTAAATTAAGCCTTCAGAAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 780
DB |||||
QY 805 TACTTAGTAAATTAAGCCTTCAGAAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 864
DB |||||
QY 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTTGGCCTCATTTGGAATGAATTCAT 840
DB |||||
QY 865 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTTGGCCTCATTTGGAATGAATTCAT 924
DB |||||
QY 841 TTGATGAACACGTACATTCAGTTTGTAGCTGACAGATTACTTTGTGAACTTCTGATTTCTCA 900
DB |||||
QY 925 TTGATGAACACGTACATTCAGTTTGTAGCTGACAGATTACTTTGTGAACTTCTGATTTCTCA 984
DB |||||
QY 901 AAGGTTTTTTCAGGAGAAAATCCTTTTGAATTTTATGAAAACATTTCTTTAGAGGAAA 960
DB |||||
QY 985 AAGGTTTTTTCAGGAGAAAATCCTTTTGAATTTTATGAAAACATTTCTTTAGAGGAAA 1044
DB |||||
QY 961 ACAATTTCTTTGAGAAACAGTTTTCAGAGTATCAGCGTTTTCAGGTTATGCGAGAAACC 1020
DB |||||
QY 1045 ACAATTTCTTTGAGAAACAGTTTTCAGAGTATCAGCGTTTTCAGGTTATGCGAGAAACC 1104
DB |||||
QY 1021 ACAGATAAGCTCTTCCACCTTTGGATGACAGATTTT 1053
DB |||||
QY 1105 ACAGATAAGCTCTTCCACCTTTGGATGACAGATTTT 1137
DB |||||

RESULT 14

US-60-449-629-12
; Sequence 12, Application US/60449629
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CANCER, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001449
; CURRENT APPLICATION NUMBER: US/60/449,629
; NUMBER OF SEQ ID NOS: 2712
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 3397
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-60-449-629-12

Query Match 99.8%; Score 1051; DB 114; Length 3397;
Best Local Similarity 99.8%; Pred. No. 3.5e-275;
Matches 1051; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATCAGAGATCATCTTCA 60
DB |||||
QY 501 ATGGGCGACCCGGAAGCGCGGAGCGCGGCTGGATCAGATCAGAGATCATCTTCA 560
DB |||||
QY 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
DB |||||
QY 561 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 620
DB |||||
QY 121 CGGTTTGTATCTTCCAAATCCAGTACCTGTATATTTGGAAAATGATATAAAGAGGACAG 180
DB |||||
QY 621 CGGTTTGTATCTTCCAAATCCAGTACCTGTATATTTGGAAAATGATATAAAGAGGACAG 680
DB |||||
QY 181 GCTTCTCTCTGACAGCAGAGAGGTTGACTTATCAAGGATCTCCCTCACTGGAAACAG 240
DB |||||
QY 681 GCTTCTCTCTGACAGCAGAGAGGTTGACTTATCAAGGATCTCCCTCACTGGAAACAG 740
DB |||||
QY 241 CTTAAAGCAGATCAGAAAGTACTTCTCATCTCTAGCCCTTTTTCGAGCCAGTGAT 300
DB |||||
QY 741 CTTAAAGCAGATCAGAAAGTACTTCTCATCTCTAGCCCTTTTTCGAGCCAGTGAT 800
DB |||||
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGAGTGAGGTTCCAGAGGCT 360
DB |||||
QY 801 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGAGTGAGGTTCCAGAGGCT 860
DB |||||
QY 361 CGCTGTTCTATGCGCTTTTCAAAATCTCATCGAGATGTTCTCATCAGAGATGACAGTTTG 420
DB |||||
QY 861 CGCTGTTCTATGCGCTTTTCAAAATCTCATCGAGATGTTCTCATCAGAGATGACAGTTTG 920
DB |||||
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATGAA 480
DB |||||
QY 921 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATGAA 980
DB |||||
QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGCCCTTCGCGATGAGATAGAGATAAGAAA 540
DB |||||
QY 981 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGCCCTTCGCGATGAGATAGAGATAAGAAA 1040
DB |||||
QY 541 TCTACTTTTGGGAAAAGAGTGTTGGCTTTGCTGCTGATAGAGAGTTTCTTCTCAGGA 600
DB |||||
QY 1041 TCTACTTTTGGGAAAAGAGTGTTGGCTTTGCTGCTGATAGAGAGTTTCTTCTCAGGA 1100
DB |||||
QY 601 TCTTTTGTGCTATATTCTGCTTAAGAGAGAGGCTTTATGCCAGACTCACTTTTTC 660
DB |||||
QY 1101 TCTTTTGTGCTATATTCTGCTTAAGAGAGAGGCTTTATGCCAGACTCACTTTTTC 1160
DB |||||
QY 661 AATGAACCTCATCAGCAGAGATGAAGACTTCTACTGTGACTTTTGTGCTGATGTTCCAA 720
DB |||||
QY 1161 AATGAACCTCATCAGCAGAGATGAAGACTTCTACTGTGACTTTTGTGCTGATGTTCCAA 1220
DB |||||
QY 721 TACTTAGTAAATTAAGCCTTCAGAAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 780
DB |||||

Db 1221 TACTTAGTAATAAGCTTTCAGAAAGAGGTCAGGAGATCATTTGATGCTGTCAAA 1280
Qy 781 ATTGACGAGGTTTTTAACAGAGCCTTGCAGAGTTGGCTCATTTGGAATGAATTCATT 840
Db 1281 ATTGACGAGGTTTTTAACAGAGCCTTGCAGAGTTGGCTCATTTGGAATGAATTCATT 1340
Qy 841 TTGATGAACAGTACATTCAGATTTGTAGCTGACAGATTAATCTTGTGAACTTGGATTTCTCA 900
Db 1341 TTGATGAACAGTACATTCAGATTTGTAGCTGACAGATTAATCTTGTGAACTTGGATTTCTCA 1400
Qy 901 AAGTTTTTCAGGACAGAAATCTTTTGAATTTATGGAAGAAATCTTTTGAAGAAAA 960
Db 1401 AAGTTTTTCAGGACAGAAATCTTTTGAATTTATGGAAGAAATCTTTTGAAGAAAA 1460
Qy 961 ACAATTTCTTTCAGAGAACGAGTTTCAGAGTATCAGCGTTTTCAGTATTCAGAGAAC 1020
Db 1461 ACAATTTCTTTCAGAGAACGAGTTTCAGAGTATCAGCGTTTTCAGTATTCAGAGAAC 1520
Qy 1021 ACAGATAACGCTCTTACCTTTGGATGACAGATTTT 1053
Db 1521 ACAGATAACGCTCTTACCTTTGGATGACAGATTTT 1553

RESULT 15

US-09-629-469A-12810
; Sequence 12810, Application US/09629469A
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: SAITO, KAORU
; APPLICANT: YAMAMOTO, JUNICHI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: WAKAMATSU, AI
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: OTSUKI, TETSUJI
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH CDNA AND THEIR USE
; FILE REFERENCE: 084335/0123
; CURRENT APPLICATION NUMBER: US/09/629,469A
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: JP 1999-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 1999-300253
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/183,322
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 19025
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12810
; LENGTH: 1601
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (40)..(1092)
US-09-629-469A-12810

Query Match 99.5%; Score 1048.2; DB 28; Length 1601;
Best Local Similarity 99.7%; Pred. No. 1.5e-274;
Matches 1050; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATGGGCGACCCGAAAGCGGAGCGGCGGCTGGATCAGATGAGATCATCTTCA 60

Db 40 ATGGGCGACCCGAAAGCGGAGCGGCGGCTGGATCAGATGAGATCATCTTCA 99
Qy 61 GACACCAAGCAAGTCAAAATAAAGTCAAAATGAAGAGCCACTCCCTAAGAAAGAGTTCTCGC 120
Db 100 GACACCAAGCAAGTCAAAATAAAGTCAAAATGAAGAGCCACTCCCTAAGAAAGAGTTCTCGC 159
Qy 121 CGGTTTGTCTCATCTTTCCAAATCCAGTACCTCTGATATTTGGAAATATGTAATAACAGGCACAG 180
Db 160 CGGTTTGTCTCATCTTTCCAAATCCAGTACCTCTGATATTTGGAAATATGTAATAACAGGCACAG 219
Qy 181 GCTTCCTCTTGGACAGCAGAGAGGTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 240
Db 220 GCTTCCTCTTGGACAGCAGAGAGGTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 279
Qy 241 CTTAAAGCAGATGAGAGAGTACTTCTCATCTCTCAGATCTTAGCCCTTTTTCAGCCAGTGAT 300
Db 280 CTTAAAGCAGATGAGAGAGTACTTCTCATCTCTCAGATCTTAGCCCTTTTTCAGCCAGTGAT 339
Qy 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTTAGTCAGGAGGTGCAAGTTCCAGAGGCT 360
Db 340 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTTAGTCAGGAGGTGCAAGTTCCAGAGGCT 399
Qy 361 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAATGTTCACTCAGAGATGTAACAGTTTG 420
Db 400 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAATGTTCACTCAGAGATGTAACAGTTTG 459
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAA 480
Db 460 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTTGAA 519
Qy 481 ACCATGCCCTATGTTAAGAAAGAGAGTATGCGCTTTCGATGAGATGAGAGAGAGAG 540
Db 520 ACCATGCCCTATGTTAAGAAAGAGAGTATGCGCTTTCGATGAGATGAGAGAGAGAG 579
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAAGAGTGTCTTCTCTCAGGA 600
Db 580 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAAGAGTGTCTTCTCTCAGGA 639
Qy 601 TCTTTTGTCTGTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGAGTCACTTTTTC 660
Db 640 TCTTTTGTCTGTATATTTCTGGCTAAAGAGAGAGAGGTCTTATGCCAGGAGTCACTTTTTC 699
Qy 661 AATGAACATCATCAGCAGAGATGAAGACTTCTCAGTGTGACTTTGCTTCCCTGATGTTCCAA 720
Db 700 AATGAACATCATCAGCAGAGATGAAGACTTCTCAGTGTGACTTTGCTTCCCTGATGTTCCAA 759
Qy 721 TACTTAGTAATAAGCTTTCAGAAAGAGGTCAGGAGAGATCATTTGTTGATGCTGTCAAA 780
Db 760 TACTTAGTAATAAGCTTTCAGAAAGAGGTCAGGAGAGATCATTTGTTGATGCTGTCAAA 819
Qy 781 ATTGACGAGAGTTTTTAAACAGAGCCTTGCAGTGGCCCTCATTTGGAATGAATTCATT 840
Db 820 ATTGACGAGAGTTTTTAAACAGAGCCTTGCAGTGGCCCTCATTTGGAATGAATTCATT 879
Qy 841 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTTGTGGAATTTGGATTTCTCA 900
Db 880 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTTGTGGAATTTGGATTTCTCA 939
Qy 901 AAGTTTTTTCAGGACAGAAATCTTTTGAATTTTATGGAAGAAATCTTTTATAGAGGAAAA 960
Db 940 AAGTTTTTTCAGGACAGAAATCTTTTGAATTTTATGGAAGAAATCTTTTATAGAGGAAAA 999
Qy 961 ACAATTTCTTTCAGAGAACGAGTTTCAGAGTATCAGCGTTTTCAGTATTCAGAGAAC 1020
Db 1000 ACAATTTCTTTCAGAGAACGAGTTTCAGAGTATCAGCGTTTTCAGTATTCAGAGAAC 1059
Qy 1021 ACAGATAACGCTCTTACCTTTGGATGACAGATTTT 1053
Db 1060 ACAGATAACGCTCTTACCTTTGGATGACAGATTTT 1092

Search completed: October 30, 2005, 05:15:34

Job time : 3217.5 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 30, 2005, 00:02:51 ; Search time 1372.5 Seconds
(without alignments)
11030.942 Million cell updates/sec

Title: US-10-698-228-12
Perfect score: 1053
Sequence: 1 atggcgaccggaagcc.....tcacctggatgcagatttt 1053

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27723106 seqs, 7188968421 residues

Total number of hits satisfying chosen parameters: 55446212

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents NA New:*

- 1: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:*
- 2: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq2:*
- 3: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:*
- 4: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq:*
- 5: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq:*
- 6: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:*
- 7: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq1:*
- 8: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq2:*
- 9: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:*
- 10: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq10:*
- 11: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq11:*
- 12: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq2:*
- 13: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq3:*
- 14: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq4:*
- 15: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq5:*
- 16: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq6:*
- 17: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq7:*
- 18: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq8:*
- 19: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq9:*
- 20: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq:*
- 21: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq2:*
- 22: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq3:*
- 23: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq4:*
- 24: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq5:*
- 25: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq6:*
- 26: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq7:*
- 27: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq8:*
- 28: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1051.4	99.8	1056	1	PCT-US05-09639-16
2	1051.4	99.8	1056	27	US-11-088-686-16
3	1051.4	99.8	4929	1	PCT-US05-22501-1545
4	1051.4	99.8	4955	1	PCT-US05-09639-44
5	1051.4	99.8	4955	1	PCT-US05-21650-13

6	1051.4	99.8	4955	27	US-11-088-686-44	Sequence 44, Appl
7	1051	99.8	1849	15	US-10-990-328A-49	Sequence 49, Appl
8	1049.8	99.7	1056	1	PCT-US05-09639-18	Sequence 18, Appl
9	1049.8	99.7	1056	27	US-11-088-686-18	Sequence 18, Appl
10	1048.2	99.5	1056	1	PCT-US05-09639-20	Sequence 20, Appl
11	1048.2	99.5	1056	27	US-11-088-686-20	Sequence 20, Appl
12	1046.6	99.4	1056	1	PCT-US05-09639-22	Sequence 22, Appl
13	1046.6	99.4	1056	1	PCT-US05-09639-24	Sequence 24, Appl
14	1046.6	99.4	1056	27	US-11-088-686-22	Sequence 22, Appl
15	1046.6	99.4	1056	27	US-11-088-686-24	Sequence 24, Appl
16	1045	99.2	1056	1	PCT-US05-09639-26	Sequence 26, Appl
17	1045	99.2	1056	1	PCT-US05-09639-28	Sequence 28, Appl
18	1045	99.2	1056	27	US-11-088-686-26	Sequence 26, Appl
19	1045	99.2	1056	27	US-11-088-686-28	Sequence 28, Appl
20	624	59.3	624	1	PCT-US05-09639-42	Sequence 42, Appl
21	624	59.3	624	27	US-11-088-686-42	Sequence 42, Appl
22	580	55.1	1170	1	PCT-US05-09639-2	Sequence 2, Appl
23	580	55.1	1170	27	US-11-088-686-2	Sequence 2, Appl
24	580	55.1	2482	11	US-10-302-689A-107255	Sequence 107255,
25	580	55.1	2500	1	PCT-US05-09639-43	Sequence 43, Appl
26	580	55.1	2500	1	PCT-US05-20435-12	Sequence 12, Appl
27	580	55.1	2500	1	PCT-US05-21650-12	Sequence 12, Appl
28	580	55.1	2500	1	PCT-US05-27243-78	Sequence 78, Appl
29	580	55.1	2500	14	US-10-960-414-154	Sequence 154, App
30	580	55.1	2500	15	US-10-940-774A-145	Sequence 145, App
31	580	55.1	2500	24	US-11-000-688A-1211	Sequence 1211, Ap
32	580	55.1	2500	24	US-11-150-888-12	Sequence 12, Appl
33	578.4	54.9	1170	1	PCT-US05-09639-4	Sequence 4, Appl
34	578.4	54.9	1170	27	US-11-088-686-4	Sequence 4, Appl
35	578.4	54.9	1649	9	US-10-553-098-583	Sequence 583, App
36	578.4	54.9	1649	13	US-10-955-054A-135	Sequence 135, App
37	578.4	54.9	2470	1	PCT-US05-13260-66	Sequence 66, Appl
38	578.4	54.9	2470	1	PCT-US05-20840-66	Sequence 66, Appl
39	578.4	54.9	2470	1	PCT-US05-22501-6487	Sequence 6487, Ap
40	578.4	54.9	2470	15	US-10-940-774A-2025	Sequence 2025, Ap
41	578.4	54.9	2909	15	US-10-990-328A-4852	Sequence 4852, Ap
42	578	54.9	3616	15	US-10-990-328A-4851	Sequence 4851, Ap
43	578	54.9	1170	1	PCT-US05-09639-6	Sequence 6, Appl
44	576.8	54.8	1170	1	PCT-US05-09639-8	Sequence 8, Appl
45	576.8	54.8	1170	1	PCT-US05-09639-8	

ALIGNMENTS

RESULT 1
PCT-US05-09639-16
; Sequence 16, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1056)
PCT-US05-09639-16

Query Match 99.8%; Score 1051.4; DB 1; Length 1056;
Best Local Similarity 99.9%; Pred. No. 3.3e-270;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGCGACCCGGAAGCCGCGGAGCCGCGGCTGCATCAGATGAGATCATCTTCA 60
|||||


```
Db 721 TACTTAGTAATAAGCCTTCAGAAAGGAGGTTCAGGAGATCATTTGTCATGCTGTCAAA 780
QY 781 ATTGACGAGAGTTTTTAACAGAGCCTTGCAGTTGGCTCATTTGGAATGAATTCGATT 840
Db 781 ATTGACGAGAGTTTTTAACAGAGCCTTGCAGTTGGCTCATTTGGAATGAATTCGATT 840
QY 841 TTGATGAACAGTACATTCAGTTGTTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
Db 841 TTGATGAACAGTACATTCAGTTGTTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
QY 901 AAGGTTTTTCAGGAGAGAAATCCTTTTGATTTTATGAGAAACATTTCTTTAGAGGAAA 960
Db 901 AAGGTTTTTCAGGAGAGAAATCCTTTTGATTTTATGAGAAACATTTCTTTAGAGGAAA 960
QY 961 ACAAAATTTCTTGAGAAACGATTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC 1020
Db 961 ACAAAATTTCTTGAGAAACGATTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC 1020
QY 1021 ACAGATAACGCTCTTCACTTGGATGCGAGATTTT 1053
Db 1021 ACAGATAACGCTCTTCACTTGGATGCGAGATTTT 1053

RESULT 3
PCT-US05-22501-1545
; Sequence 1545, Application PC/TUS0522501
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals
; TITLE OF INVENTION: Determining Cancer-Linked Gene and Therapeutic Targets Using
; FILE OF INVENTION: Molecular Cytogenetic Methods
; CURRENT APPLICATION NUMBER: PCT/US05/22501
; CURRENT FILING DATE: 2005-07-07
; PRIOR APPLICATION NUMBER: 60/581,699
; PRIOR FILING DATE: 2004-06-23
; NUMBER OF SEQ ID NOS: 7840
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1545
; LENGTH: 4929
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: cDNA sequence
PCT-US05-22501-1545

Query Match 99.8%; Score 1051.4; DB 1; Length 4929;
Best Local Similarity 99.9%; Pred. No. 5.1e-270;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGGGACCCGGAAAGCGCGGAGCGCGGCTGGATCAGAGTGAGAGATCATCTTCA 60
Db 245 ATGGGGGACCCGGAAAGCGCGGAGCGCGGCTGGATCAGAGTGAGAGATCATCTTCA 304
QY 61 GACACCAAGGAGTGAATAAGTCAATGAAGGCACTCTTAAGAAAGGTTCTCCG 120
Db 305 GACACCAAGGAGTGAATAAGTCAATGAAGGCACTCTTAAGAAAGGTTCTCCG 364
QY 121 CGGTTTGTTCATCTTTCCAAATCCAGTACCTGATATTTGGAAATGTATATAACAGGCACAG 180
Db 365 CGGTTTGTTCATCTTTCCAAATCCAGTACCTGATATTTGGAAATGTATATAACAGGCACAG 424
QY 181 GCTTCCTTCTGACAGCAGAGAGGTTGACTTATCAAAAGGATCTCCCTCACTGGAAACAAG 240
Db 425 GCTTCCTTCTGACAGCAGAGAGGTTGACTTATCAAAAGGATCTCCCTCACTGGAAACAAG 484
QY 241 CTTAAGCAGATGAGAACTTCTATCTCTACATCTTAGCCCTTTTTCAGCCAGTGTAT 300
Db 485 CTTAAGCAGATGAGAACTTCTATCTCTACATCTTAGCCCTTTTTCAGCCAGTGTAT 544
QY 301 GGAATTTGTAATAATTTTGGTGAGCGCTTTAGTCAGGAGGTGTCAGGTTCCAGAGGCT 360
Db 545 GGAATTTGTAATAATTTTGGTGAGCGCTTTAGTCAGGAGGTGTCAGGTTCCAGAGGCT 604
```

```
QY 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAGAAATGTTCTACTCAGAGATGTACAGTTTG 420
Db 605 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAGAAATGTTCTACTCAGAGATGTACAGTTTG 664
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTTTATTTAAATGCAATGAA 480
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTTTATTTAAATGCAATGAA 724
QY 481 ACCATGCCCTATGTTAAGAAAAGAGATTTGGCCCTTCCGATGGATGAGCAGATAGAAA 540
Db 725 ACCATGCCCTATGTTAAGAAAAGAGATTTGGCCCTTCCGATGGATGAGCAGATAGAAA 784
QY 541 TCTACTTTTGGGAAAAGAGTGGTGGCTTTGCTCTGTGAGAAAGAGTTTTTCTTCTCAGAA 600
Db 785 TCTACTTTTGGGAAAAGAGTGGTGGCTTTGCTCTGTGAGAAAGAGTTTTTCTTCTCAGAA 844
QY 601 TCTTTTGTCTCTATTTCTGGCTAAAGAAGAGAGGTCTTATGCCAGGACTCACATTTTTC 660
Db 845 TCTTTTGTCTCTATTTCTGGCTAAAGAAGAGAGGTCTTATGCCAGGACTCACATTTTTC 904
QY 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGGCTGATGTTC 720
Db 905 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGGCTGATGTTC 964
QY 721 TACTTAGTAATAAGCCTTCAGAGAAAGGTCAGGAGAGATCATTTGTGATGCTGTCAAA 780
Db 965 TACTTAGTAATAAGCCTTCAGAGAAAGGTCAGGAGAGATCATTTGTGATGCTGTCAAA 1024
QY 781 ATTGAGCAGGAGTTTTTAAACAGAGCCCTGCGAGTTCGCCCTCATTTGGAATGAATTCGATT 840
Db 1025 ATTGAGCAGGAGTTTTTAAACAGAGCCCTGCGAGTTCGCCCTCATTTGGAATGAATTCGATT 1084
QY 841 TTGATGAACAGTACATTCAGTTGTTGATCTGACAGATTTACTTTGGAACTTTGGATTTCTCA 900
Db 1085 TTGATGAACAGTACATTCAGTTGTTGATCTGACAGATTTACTTTGGAACTTTGGATTTCTCA 1144
QY 901 AAGTTTTTTCAGGAGAGAAATCCTTTTGATTTTATGAAACATTTCTTTAGAAAGGAAA 960
Db 1145 AAGTTTTTTCAGGAGAGAAATCCTTTTGATTTTATGAAACATTTCTTTAGAAAGGAAA 1204
QY 961 ACAAAATTTCTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTTGCAGTTATGCGAGAAACC 1020
Db 1205 ACAAAATTTCTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTTGCAGTTATGCGAGAAACC 1264
QY 1021 ACAGATAACGCTCTTCACTTGGATGCGAGATTTT 1053
Db 1265 ACAGATAACGCTCTTCACTTGGATGCGAGATTTT 1297
```

```
RESULT 4
PCT-US05-09639-44
; Sequence 44, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US05-09639-44
```

```
Query Match 99.8%; Score 1051.4; DB 1; Length 4955;
Best Local Similarity 99.9%; Pred. No. 5.1e-270;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```



```
QY 1 ATGGCGACCCCGAAGCGCGGAGCGCGCGGCTGGATCAGATGAGAGATCATCTTCA 60
Db 245 ATGGCGACCCCGAAGCGCGGAGCGCGCGGCTGGATCAGATGAGAGATCATCTTCA 304
QY 61 GACACCAAGCAAGTGAATAAAGTCAATGAAGAGCGCACTCTTAAGAAGAGTTCTCGC 120
Db 305 GACACCAAGCAAGTGAATAAAGTCAATGAAGAGCGCACTCTTAAGAAGAGTTCTCGC 364
QY 121 CGTTTGTGTCATCTTCCCAATCAGTACCTGATATTTGGAAAATGTTATAAAGAGGACAG 180
Db 365 CGTTTGTGTCATCTTCCCAATCAGTACCTGATATTTGGAAAATGTTATAAAGAGGACAG 424
QY 181 GCTTCTCTTGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 240
Db 425 GCTTCTCTTGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 484
QY 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTCTAGGCTTTTTCGAGCCAGTGAT 300
Db 485 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTCTAGGCTTTTTCGAGCCAGTGAT 544
QY 301 GGAATTTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGGAGTGCGAGGCT 360
Db 545 GGAATTTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGGAGTGCGAGGCT 604
QY 361 CGCTGTTTCTATGCGCTTTCAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTG 420
Db 605 CGCTGTTTCTATGCGCTTTCAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTG 664
QY 421 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTGAA 480
Db 665 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTGAA 724
QY 481 ACCATGCCCTATCTTAAGAAAAGCAGATTTGGCCCTTGGATGATAGAGATAGAAA 540
Db 725 ACCATGCCCTATCTTAAGAAAAGCAGATTTGGCCCTTGGATGATAGAGATAGAAA 784
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCCCTTTGCTGCTAGAGAGATTTTCTTCTCAGGA 600
Db 785 TCTACTTTTGGGAAAGAGTGGTGGCCCTTTGCTGCTAGAGAGATTTTCTTCTCAGGA 844
QY 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGCTTATGCCAGGACTCACTTTTTC 660
Db 845 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGCTTATGCCAGGACTCACTTTTTC 904
QY 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA 720
Db 905 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA 964
QY 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGCTCAGGAGAGATCATTTGTTGATGCTGTCAA 780
```

RESULT 5

```
PCT-US05-21650-13
; Sequence 13, Application PC/TUS0521650
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: MIGFS AS MODIFIERS OF THE IGF PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX05-017
; CURRENT APPLICATION NUMBER: PCT/US05/21650
; CURRENT FILING DATE: 2005-06-27
; PRIOR APPLICATION NUMBER: 60/581,689
; PRIOR FILING DATE: 2004-06-21
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US05-21650-13
```

```
Query Match 99.8%; Score 1051.4; DB 1; Length 4955;
Best Local Similarity 99.9%; Pred. No. 5.1e-270;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 ATGGCGACCCCGAAGCGCGGAGCGCGCGGCTGGATCAGATGAGAGATCATCTTCA 60
Db 245 ATGGCGACCCCGAAGCGCGGAGCGCGCGGCTGGATCAGATGAGAGATCATCTTCA 304
QY 61 GACACCAAGCAAGTGAATAAAGTCAATGAAGAGCGCACTCTTAAGAAGAGTTCTCGC 120
Db 305 GACACCAAGCAAGTGAATAAAGTCAATGAAGAGCGCACTCTTAAGAAGAGTTCTCGC 364
QY 121 CGTTTGTGTCATCTTCCCAATCAGTACCTGATATTTGGAAAATGTTATAAAGAGGACAG 180
Db 365 CGTTTGTGTCATCTTCCCAATCAGTACCTGATATTTGGAAAATGTTATAAAGAGGACAG 424
QY 181 GCTTCTCTTGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 240
Db 425 GCTTCTCTTGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAACAAG 484
QY 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTCTAGGCTTTTTCGAGCCAGTGAT 300
Db 485 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTCTAGGCTTTTTCGAGCCAGTGAT 544
QY 301 GGAATTTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGGAGTGCGAGGCT 360
Db 545 GGAATTTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGGAGTGCGAGGCT 604
QY 361 CGCTGTTTCTATGCGCTTTCAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTG 420
Db 605 CGCTGTTTCTATGCGCTTTCAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTG 664
QY 421 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTGAA 480
Db 665 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTGAA 724
QY 481 ACCATGCCCTATCTTAAGAAAAGCAGATTTGGCCCTTGGATGATAGAGATAGAAA 540
Db 725 ACCATGCCCTATCTTAAGAAAAGCAGATTTGGCCCTTGGATGATAGAGATAGAAA 784
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCCCTTTGCTGCTAGAGAGATTTTCTTCTCAGGA 600
Db 785 TCTACTTTTGGGAAAGAGTGGTGGCCCTTTGCTGCTAGAGAGATTTTCTTCTCAGGA 844
QY 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGCTTATGCCAGGACTCACTTTTTC 660
Db 845 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGCTTATGCCAGGACTCACTTTTTC 904
QY 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA 720
Db 905 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA 964
QY 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGCTCAGGAGAGATCATTTGTTGATGCTGTCAA 780
```

Db 965 TACTTAGTAATAAAGCCTTCAGAAAGGAGGTGAGGAGATCATTTGATGCTGTCAAA 1024
QY 781 ATTGAGCAGGAGTCTTTTAAAGAGGCTTGGCAGTTGGCTCATTTGGAATGAATTCATT 840
Db 1025 ATTGAGCAGGAGTCTTTTAAAGAGGCTTGGCAGTTGGCTCATTTGGAATGAATTCATT 1084
QY 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTACTTTGTGAACTTGGATTCTCA 900
Db 1085 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTACTTTGTGAACTTGGATTCTCA 1144
QY 901 AAGGTTTTTCAGGAGAGAAATCTTTTGAATTTTATGGAAGAAACATTTCTTTAGAGGAAAA 960
Db 1145 AAGGTTTTTCAGGAGAGAAATCTTTTGAATTTTATGGAAGAAACATTTCTTTAGAGGAAAA 1204
QY 961 ACAAAATTTCTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTGTATGCGAGAAACC 1020
Db 1205 ACAAAATTTCTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTGTATGCGAGAAACC 1264
QY 1021 ACAGATAACGCTTTCACCTTTGGATGAGATTTT 1053
Db 1265 ACAGATAACGCTTTCACCTTTGGATGAGATTTT 1297

RESULT 6

US-11-088-686-44

; Sequence 44, Application US/11088686

; GENERAL INFORMATION:

; APPLICANT: Yen, Yun

; TITLE OF INVENTION: DRUG SCREENING

; FILE REFERENCE: 14037-004001

; CURRENT APPLICATION NUMBER: US/11/088,686

; CURRENT FILING DATE: 2005-03-23

; PRIOR APPLICATION NUMBER: US 60/556,836

; PRIOR FILING DATE: 2004-03-25

; NUMBER OF SEQ ID NOS: 59

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 44

; LENGTH: 4955

; TYPE: DNA

; ORGANISM: Homo sapiens

US-11-088-686-44

Query Match 99.8%; Score 1051.4; DB 27; Length 4955;

Best Local Similarity 99.9%; Pred. No. 5.1e-270;

Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCCGAAGCGCGGAGCGGCTGGATCAGATGAGAGATCATCTTCA 60
Db 245 ATGGGCGACCCCGAAGCGCGGAGCGGCTGGATCAGATGAGAGATCATCTTCA 304
QY 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 305 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 364
QY 121 CGGTTTGTCTCTTCCAAATCCAGTACCTGTATTTGGAAATGTATTAACAGGCACAG 180
Db 365 CGGTTTGTCTCTTCCAAATCCAGTACCTGTATTTGGAAATGTATTAACAGGCACAG 424
QY 181 GCTTCTCTTCTGACAGCAGAGAGGTTGACTTATCAAGGATCTCCCTCAGTGGAAACAG 240
Db 425 GCTTCTCTTCTGACAGCAGAGAGGTTGACTTATCAAGGATCTCCCTCAGTGGAAACAG 484
QY 241 CTTAAAGCAGATGAGAAGTACTTCTCTCAATCTTAGCCCTTTTTCAGCCAGTGTAT 300
Db 485 CTTAAAGCAGATGAGAAGTACTTCTCTCAATCTTAGCCCTTTTTCAGCCAGTGTAT 544
QY 301 GGAATTTGAATGAATAATTTGTGAGCGCTTTAGTCAGGAGGTGAGGTTCCAGAGGT 360
Db 545 GGAATTTGAATGAATAATTTGTGAGCGCTTTAGTCAGGAGGTGAGGTTCCAGAGGT 604
QY 361 CGCTGTTCTATGGCTTTTCAAAATCTTCATCAGAGATGTTCACTCAGAGATGACAGTTTG 420
Db 605 CGCTGTTCTATGGCTTTTCAAAATCTTCATCAGAGATGTTCACTCAGAGATGACAGTTTG 664

QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTATTTAAATGCAATTGAA 480
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTATTTAAATGCAATTGAA 724
QY 481 ACCATGCCCTATGTTTAAGAAAAAGCAGATTGGGCCCTTGGATGGATAGCAGATAGAAA 540
Db 725 ACCATGCCCTATGTTTAAGAAAAAGCAGATTGGGCCCTTGGATGGATAGCAGATAGAAA 784
QY 541 TCTACTTTTGGGAAAGAGTGGGCCCTTGGCTCTCTGTAGAGAGGATTTTCTTCTCAGGA 600
Db 785 TCTACTTTTGGGAAAGAGTGGGCCCTTGGCTCTCTGTAGAGAGGATTTTCTTCTCAGGA 844
QY 601 TCTTTTGTCTATATTTCTGCTTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
Db 845 TCTTTTGTCTATATTTCTGCTTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTTC 904
QY 661 AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGTCTGATGTTCCAA 720
Db 905 AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGTCTGATGTTCCAA 964
QY 721 TACTTAGTAATAAGCCTTCAGAAAGAGGTGAGGAGATCATTTGATGCTGTCTCAAA 780
Db 965 TACTTAGTAATAAGCCTTCAGAAAGAGGTGAGGAGATCATTTGATGCTGTCTCAAA 1024
QY 781 ATTGAGCAGGAGTTTAAAGAGGCTTGGCAGTTGGCCTCATTTGGAATGAATTCATT 840
Db 1025 ATTGAGCAGGAGTTTAAAGAGGCTTGGCAGTTGGCCTCATTTGGAATGAATTCATT 1084
QY 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTACTTTGGAACCTTTGGATTCTCA 900
Db 1085 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTACTTTGGAACCTTTGGATTCTCA 1144
QY 901 AAGGTTTTTCAGGAGAGAAATCTTTTGAATTTTATGGAAGAAACATTTCTTTAGAGGAAAA 960
Db 1145 AAGGTTTTTCAGGAGAGAAATCTTTTGAATTTTATGGAAGAAACATTTCTTTAGAGGAAAA 1204
QY 961 ACAAAATTTCTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTGTATGCGAGAAACC 1020
Db 1205 ACAAAATTTCTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTGTATGCGAGAAACC 1264
QY 1021 ACAGATAACGCTTTCACCTTTGGATGAGATTTT 1053
Db 1265 ACAGATAACGCTTTCACCTTTGGATGAGATTTT 1297

RESULT 7

US-10-990-328A-49

; Sequence 49, Application US/10990328A

; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele

; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES

; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND

; TITLE OF INVENTION: USES THEREOF

; FILE REFERENCE: CL001495

; CURRENT APPLICATION NUMBER: US/10/990,328A

; CURRENT FILING DATE: 2004-11-17

; NUMBER OF SEQ ID NOS: 558824

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 49

; LENGTH: 1849

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-990-328A-49

Query Match 99.8%; Score 1051; DB 15; Length 1849;

Best Local Similarity 99.8%; Pred. No. 4.9e-270;

Matches 1051; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCCGAAGCGCGGAGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 85 ATGGGCGACCCCGAAGCGCGGAGCGGCTGGATCAGGATGAGAGATCATCTTCA 144

```
QY 61 GACACAAAGGAAATGAATGAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
Db 145 GACACAAAGGAAATGAATGAAGTCAAAATGAAGAGCCACTCTAAGAAAGAGTTCTCGC 204
QY 121 CGGTTTGTCTATCTTCCATCCAGTACCTGATATTTGGAATGTAATAAGAGGACAG 180
Db 205 CGGTTTGTCTATCTTCCATCCAGTACCTGATATTTGGAATGTAATAAGAGGACAG 264
QY 181 GCTTCTCTCTGACAGCAGAGAGGTTGACTTATCAAGGATCTCCTCACTGGAACAAG 240
Db 265 GCTTCTCTCTGACAGCAGAGAGGTCGACTTATCAAGGATCTCCTCACTGGAACAAG 324
QY 241 CTTAAAGCAGATGAGAAGTCTTCACTCTCAATCTTAGCCCTTTTTCAGCCAGTGAT 300
Db 325 CTTAAAGCAGATGAGAAGTCTTCACTCTCAATCTTAGCCCTTTTTCAGCCAGTGAT 384
QY 301 GGAATTTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
Db 385 GGAATTTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 444
QY 361 CGCTGTCTCTATGGCTTTCAAAATCTCATCGAATGTTCACTCAGAGATGTACAGTTTG 420
Db 445 CGCTGTCTCTATGGCTTTCAAAATCTCATCGAATGTTCACTCAGAGATGTACAGTTTG 504
QY 421 CTGATAGACACTTACATGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 505 CTGATAGACACTTACATGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 564
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGGATGGATAGCAGATAGAAA 540
Db 565 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGGATGGATAGCAGATAGAAA 624
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAAGAGGAGTTTCTTCTCAGGA 600
Db 625 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAAGAGGAGTTTCTTCTCAGGA 684
QY 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGTTTATGCGCAGGACTCACTTTTTC 660
Db 685 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGTTTATGCGCAGGACTCACTTTTTC 744
QY 661 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 745 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 804
QY 721 TACTTAGTAATAAGCTTTCAGAGAAAGGTCAGGAGATCATTTGTCATGCTGTCAAA 780
Db 805 TACTTAGTAATAAGCTTTCAGAGAAAGGTCAGGAGATCATTTGTCATGCTGTCAAA 864
QY 781 ATTGAGCAGGAGTTTAAACAGAAAGCTTTGCCAGTTGGCTCATTTGGAATGAAATTCATT 840
Db 865 ATTGAGCAGGAGTTTAAACAGAAAGCTTTGCCAGTTGGCTCATTTGGAATGAAATTCATT 924
QY 841 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTAATTGTGGAACTTGGAAATCTCA 900
Db 925 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTAATTGTGGAACTTGGAAATCTCA 984
QY 901 AAGGTTTTTTCAGCAGAAATCTTTTGTATTTTATGGAACAATTTCTTTAGAGGAAA 960
Db 985 AAGGTTTTTTCAGCAGAAATCTTTTGTATTTTATGGAACAATTTCTTTAGAGGAAA 1044
QY 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC 1020
Db 1045 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC 1104
QY 1021 ACAGATAACGCTCTTCACTCTGGATGAGATTTT 1053
Db 1105 ACAGATAACGCTCTTCACTCTGGATGAGATTTT 1137
```

RESULT 8
PCT-US05-09639-18
; Sequence 18, Application PC/TUS0509639
; GENERAL INFORMATION:

```
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US05-09639-18

Query Match 99.7%; Score 1049.8; DB 1; Length 1056;
Best Local Similarity 99.8%; Pred. No. 8.8e-270; Indels 0; Gaps 0;
Matches 1051; Conservative 0; Mismatches 2;

QY 1 ATGGGCGACCCGGAAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 ATGGGCGACCCGGAAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60,
QY 61 GACACAAAGGAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 61 GACACAAAGGAGTGAATTAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
QY 121 CGGTTTGTCTATCTTCCATCCAGTACCTTATCAAAAGGATCTCCCTCACTGGAAACAAG 180
Db 121 CGGTTTGTCTATCTTCCATCCAGTACCTTATCAAAAGGATCTCTAATAACAGGCAAG 180
QY 181 GCTTCTCTTGGACAGCAGAGAGGTTGACTTATCAAAAGGATCTCCCTCACTGGAAACAAG 240
Db 181 GCTTCTCTTGGACAGCAGAGAGGTCGACTTATCAAAAGGATCTCCCTCACTGGAAACAAG 240
QY 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCAATCTTAGCCTTTTTCAGCCAGTGAT 300
Db 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCAATCTTAGCCTTTTTCAGCCAGTGAT 300
QY 301 GGAATTTGTAATAAGGTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCCAGAGGCT 360
Db 301 GGAATTTGTAATAAGGTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCCAGAGGCT 360
QY 361 CGCTGTCTTATGCTTTCAAAATTTCTCATCAGAAATGTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTCTTATGCTTTCAAAATTTCTCATCAGAAATGTTCACTCAGAGATGTACAGTTTG 420
QY 421 CTGATAGACACTTATCATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 421 CTGATAGACACTTATCATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGGATGGATAGCAGATAGAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGGATGGATAGCAGATAGAAA 540
QY 541 TCTACTTTTGGGAAAGAGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGGCTTTGCTGCTGTAGAGGAGTTTCTTCTCAGGA 600
QY 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGTCCTTATGCGCAGGACTCACTTTTTC 660
Db 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGTCCTTATGCGCAGGACTCACTTTTTC 660
QY 661 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTGCTGCTGATGTTTCCAA 720
Db 661 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTGCTGCTGATGTTTCCAA 720
QY 721 TACTTAGTAATAAGCTTTCAGAGAAAGGTCAGGAGATCATTTGTCATGCTGTCAAA 780
Db 721 TACTTAGTAATAAGCTTTCAGAGAAAGGTCAGGAGATCATTTGTCATGCTGTCAAA 780
QY 781 ATTGAGCAGGAGTTTAAACAGAAAGCTTTGCCAGTTGGCTCATTTGGAATGAAATTCATT 840
```

Db 781 ATTGACGAGGATTTTAAACAGAGCCTTCCAGTGTGGCCTCATTTGAATGAATTCATT 840
Qy 841 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTACTCTGGAACCTTGGATTCTCA 900
Db 841 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTACTCTGGAACCTTGGATTCTCA 900
Qy 901 AAGTTTTCAGGCAGAGAAATCCTTTGATTTTATGGAACCAATTTCTTTAGAGGAAAA 960
Db 901 AAGTTTTCAGGCAGAGAAATCCTTTGATTTTATGGAACCAATTTCTTTAGAGGAAAA 960
Qy 961 ACAAATTTCTTTGAGAAACAGAGTTTCAGATATCAGCGTTTTCAGATTATGGCAGAAACC 1020
Db 961 ACAAATTTCTTTGAGAAACAGAGTTTCAGATATCAGCGTTTTCAGATTATGGCAGAAACC 1020
Qy 1021 ACAGATAACCTCTTACCTTGGATGAGATTTT 1053
Db 1021 ACAGATAACCTCTTACCTTGGATGAGATTTT 1053

RESULT 9
US-11-088-686-18
; Sequence 18, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-088-686-18

Query Match 99.7%; Score 1049.8; DB 27; Length 1056;
Best Local Similarity 99.8%; Pred. No. 8.8e-270;
Matches 1051; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGGGGACCCGGAAGCGCGGAAGCGCGGCTGGATCAGATGAGATCATCTTCA 60
Db 1 ATGGGGACCCGGAAGCGCGGAAGCGCGGCTGGATCAGATGAGATCATCTTCA 60
Qy 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Qy 121 CGGTTTGTCACTTTCCAAATCCAGTACCTGTATTTGGAATGATTAATAACAGGCACAG 180
Db 121 CGGTTTGTCACTTTCCAAATCCAGTACCTGTATTTGGAATGATTAATAACAGGCACAG 180
Qy 181 GCTTCCTCTGACACAGAGAGGTTGACTTATCAAGAGTCTCCCTCACTGGACAG 240
Db 181 GCTTCCTCTGACACAGAGAGGTTGACTTATCAAGAGTCTCCCTCACTGGACAG 240
Qy 241 CTTAAGCAGATGAGAAGTACTTCTCATCTCTAGCCTTTTTCAGCAGCAGTAT 300
Db 241 CTTAAGCAGATGAGAAGTACTTCTCATCTCTAGCCTTTTTCAGCAGCAGTAT 300
Qy 301 GGAATTTGAATGAATTTGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
Db 301 GGAATTTGAATGAATTTGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
Qy 361 CGCTGTTCTATGGCTTTCAATTTCTCATCGAGAATGTTCACTCAGAGATGATACAGTTTG 420
Db 361 CGCTGTTCTATGGCTTTCAATTTCTCATCGAGAATGTTCACTCAGAGATGATACAGTTTG 420
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATTTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATTTGAA 480

Qy 481 ACCATGCCCTATGTTTAAAGAAAAACGATTTGGGCTTCCGATGATAGATAGAGAAAA 540
Db 481 ACCATGCCCTATGTTTAAAGAAAAACGATTTGGGCTTCCGATGATAGATAGAGAAAA 540
Qy 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAAAGAGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAAAGAGTTTCTTCTCAGGA 600
Qy 601 TCTTTTCTGCTATATTTCTGCTAAAGAGAGAGTCTTATGCCAGGACTCATTTTCTCC 660
Db 601 TCTTTTCTGCTATATTTCTGCTAAAGAGAGAGTCTTATGCCAGGACTCATTTTCTCC 660
Qy 661 AATGAACCTCATCAGCAGAGATGAAGAGCTTCACTGCTGACTTTTGTCTGCTGATGTTCCAA 720
Db 661 AATGAACCTCATCAGCAGAGATGAAGAGCTTCACTGCTGACTTTTGTCTGCTGATGTTCCAA 720
Qy 721 TACTTAGTAATAAGCCTTTCAGAAAGAGGCTCAGGAGATCATTTGTGATGCTGTCAA 780
Db 721 TACTTAGTAATAAGCCTTTCAGAAAGAGGCTCAGGAGATCATTTGTGATGCTGTCAA 780
Qy 781 ATTGACGAGGATTTTAAACAGAGCCTTGGCAGTTGGCCTCATTTGGAATGAATTCATT 840
Db 781 ATTGACGAGGATTTTAAACAGAGCCTTGGCAGTTGGCCTCATTTGGAATGAATTCATT 840
Qy 841 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTTACTTGTGAACTTGGATTCTCA 900
Db 841 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTTACTTGTGAACTTGGATTCTCA 900
Qy 901 AAGTTTTCAGGCAGAGAAATCCTTTGATTTTATGGAACCAATTTCTTTAGAGGAAAA 960
Db 901 AAGTTTTCAGGCAGAGAAATCCTTTGATTTTATGGAACCAATTTCTTTAGAGGAAAA 960
Qy 961 ACAAATTTCTTTGAGAAACAGAGTTTCAGAGTATCAGCGTTTTCAGATTATGGCAGAAACC 1020
Db 961 ACAAATTTCTTTGAGAAACAGAGTTTCAGAGTATCAGCGTTTTCAGATTATGGCAGAAACC 1020
Qy 1021 ACAGATAACCTCTTACCTTGGATGAGATTTT 1053
Db 1021 ACAGATAACCTCTTACCTTGGATGAGATTTT 1053

RESULT 10
PCT-US05-09639-20
; Sequence 20, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US05-09639-20

Query Match 99.5%; Score 1048.2; DB 1; Length 1056;
Best Local Similarity 99.7%; Pred. No. 2.4e-269;
Matches 1050; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATGGGGACCCGGAAGCGCGGAGCGGCTGGATCAGATGAGATCATCTTCA 60
Db 1 ATGGGGACCCGGAAGCGCGGAGCGGCTGGATCAGATGAGATCATCTTCA 60
Qy 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCTCTTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCTCTTAAGAAAGAGTTCTCGC 120

```
121 CGTTTGTCAATCTTCCAAATCCAGTACCCTGATATTTGGAAATGATATAAAGGACACAG 180
Db |||||||
121 CGTTTGTCAATCTTCCAAATCCAGTACCCTGATATTTGGAAATGATATAAAGGACACAG 180
Qy |||||||
181 GCTTCTTCTGACAGCAGAGGTTGACTTATCAAGGATCTCCTCACTGGAACAAG 240
Db |||||||
181 GCTTCTTCTGACAGCAGAGGTTGACTTATCAAGGATCTCCTCACTGGAACAAG 240
Qy |||||||
241 CTTAAAGCAGATGAGAGTACTCTCTCATCTTACATCTTACGCTTTTTCAGCCAGTGAT 300
Db |||||||
241 CTTAAAGCAGATGAGAGTACTCTCTCATCTTACATCTTACGCTTTTTCAGCCAGTGAT 300
Qy |||||||
301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTGACAGTTCCAGAGGCT 360
Db |||||||
301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTGACAGTTCCAGAGGCT 360
Qy |||||||
361 CGCTGTTCTATGGCTTTTCAAAATCTCATCGAAGATTTTCACTCAGAGATGTACAGTTTG 420
Db |||||||
361 CGCTGTTCTATGGGCTTTTCAAAATCTCATCGAAGATTTTCACTCAGAGATGTACAGTTTG 420
Qy |||||||
421 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db |||||||
421 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Qy |||||||
481 ACCATGCCCTATGTTAAGAAAGGATTTGGGCTTTGCGATGGATAGCAGATAGAAA 540
Db |||||||
481 ACCATGCCCTATGTTAAGAAAGGATTTGGGCTTTGCGATGGATAGCAGATAGAAA 540
Qy |||||||
541 TCTACTTTTGGGAAAGAGTGGCTTTTCAATCTCATCGAAGATTTTCACTCAGAGATGTACAGTTTG 600
Db |||||||
541 TCTACTTTTGGGAAAGAGTGGCTTTTCAATCTCATCGAAGATTTTCACTCAGAGATGTACAGTTTG 600
Qy |||||||
601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGGATTTTATGCGAGACTCACTTTTTC 660
Db |||||||
601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGGATTTTATGCGAGACTCACTTTTTC 660
Qy |||||||
661 AATGAACATCTACAGCAGAGATGAAGGATTTGCGCTTTGCTGTGCTGATGTTCCAA 720
Db |||||||
661 AATGAACATCTACAGCAGAGATGAAGGATTTGCGCTTTGCTGTGCTGATGTTCCAA 720
Qy |||||||
721 TACTTAGTAATAAGCCTTCAGAAAGGCTCAGAGATCATTTGTTGATGCTGTCAAA 780
Db |||||||
721 TACTTAGTAATAAGCCTTCAGAAAGGCTCAGAGATCATTTGTTGATGCTGTCAAA 780
Qy |||||||
781 ATTGAGCAGAGTTTAAAGAGGCTTGGCAGTTGGCCTCATTTGGAATGAATTGCATT 840
Db |||||||
781 ATTGAGCAGAGTTTAAAGAGGCTTGGCAGTTGGCCTCATTTGGAATGAATTGCATT 840
Qy |||||||
841 TTGATGAAACAGTACATTTGATGTTGTAGCTGACAGATTTACTTGGAACTTGGATTCTCA 900
Db |||||||
841 TTGATGAAACAGTACATTTGATGTTGTAGCTGACAGATTTACTTGGAACTTGGATTCTCA 900
Qy |||||||
901 AAGGTTTTTTCAGGAGAAATCTTTGATTTATGGAACAATTTCTTTTAGAAGGAAA 960
Db |||||||
901 AAGGTTTTTTCAGGAGAAATCTTTGATTTATGGAACAATTTCTTTTAGAAGGAAA 960
Qy |||||||
961 ACAAAATTTCTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGGATTTGCGAGAAACC 1020
Db |||||||
961 ACAAAATTTCTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGGATTTGCGAGAAACC 1020
Qy |||||||
1021 ACAGATAACGCTTTCACCTTTGGATGACAGATTTT 1053
Db |||||||
1021 ACAGATAACGCTTTCACCTTTGGATGACAGATTTT 1053
```

RESULT 11

```
US-11-088-686-20
; Sequence 20, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088, 686
```

```
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-088-686-20

Query Match      99.5%; Score 1048.2; DB 27; Length 1056;
Best Local Similarity 99.7%; Pred. No. 2.4e-269;
Matches 1050; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATCGGCGACCCGGAAGCGCGGCGTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 ATCGGCGACCCGGAAGCGCGGCGTGGATCAGGATGAGAGATCATCTTCA 60
Qy 61 GACACCAACGAAAGTGAATATAAGTCAAATCAAGAGCCACTCTCTAAGAAAGAGTTCTGCG 120
Db 61 GACACCAACGAAAGTGAATATAAGTCAAATCAAGAGCCACTCTCTAAGAAAGAGTTCTGCG 120
Qy 121 CGGTTTGTCACTTTTCAATCCAGTACCTCTGATATTTGGAAATGTATAAACAGGCAAG 180
Db 121 CGGTTTGTCACTTTTCAATCCAGTACCTCTGATATTTGGAAATGTATAAACAGGCAAG 180
Qy 181 GCTTCTTCTGGACAGCAGAGAGGTGACTTATCAAAAGGATCTCCCTCACTGGAACAAG 240
Db 181 GCTTCTTCTGGACAGCAGAGAGGTGACTTATCAAAAGGATCTCCCTCACTGGAACAAG 240
Qy 241 CTTAAAGCAGATCAGAGTACTTCTCATCTTCAATCTTAGCCTTTTTCAGCCAGTGAT 300
Db 241 CTTAAAGCAGATCAGAGTACTTCTCATCTTCAATCTTAGCCTTTTTCAGCCAGTGAT 300
Qy 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTGACAGGCT 360
Db 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGTGACAGGCT 360
Qy 361 CGCTGTTCTATGGCTTTTCAATCTCATCGAAGATTTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTCTATGGGCTTTTCAAAATCTCATCGAAGATTTTCACTCAGAGATGTACAGTTTG 420
Qy 421 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 421 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Qy 481 ACCATGCCCTATGTTAAGAAAGGATTTGGGCTTTGCGATGGATAGCAGATAGAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAGGATTTGGGCTTTGCGATGGATAGCAGATAGAAA 540
Qy 541 TCTACTTTTGGGAAAGAGTGGCTTTTCAATCTCATCGAAGATTTTCACTCAGAGATGTACAGTTTG 600
Db 541 TCTACTTTTGGGAAAGAGTGGCTTTTCAATCTCATCGAAGATTTTCACTCAGAGATGTACAGTTTG 600
Qy 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGGATTTTATGCGAGACTCACTTTTTC 660
Db 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGGATTTTATGCGAGACTCACTTTTTC 660
Qy 661 AATGAACATCTACAGCAGAGATGAAGGATTTGCGCTTTGCTGTGCTGATGTTCCAA 720
Db 661 AATGAACATCTACAGCAGAGATGAAGGATTTGCGCTTTGCTGTGCTGATGTTCCAA 720
Qy 721 TACTTAGTAATAAGCCTTCAGAAAGGCTCAGAGATCATTTGTTGATGCTGTCAAA 780
Db 721 TACTTAGTAATAAGCCTTCAGAAAGGCTCAGAGATCATTTGTTGATGCTGTCAAA 780
Qy 781 ATTGAGCAGAGTTTAAAGAGGCTTGGCAGTTGGCCTCATTTGGAATGAATTGCATT 840
Db 781 ATTGAGCAGAGTTTAAAGAGGCTTGGCAGTTGGCCTCATTTGGAATGAATTGCATT 840
Qy 841 TTGATGAAACAGTACATTTGATGTTGTAGCTGACAGATTTACTTGGAACTTGGATTCTCA 900
Db |||||||
```

```
Db      841 TTGATGAACAGTACATGAGTTGTAGCTGACAGATTACTTGTGAACTGGATTCTCA 900
Qy      901 AAGGTTTTTTCAGGCAGAGAAATCCTTTTGTATTTATGAAACATTTCTTTAGAAAGAAA 960
Db      901 AAGGTTTTTTCAGGCAGAGAAATCCTTTTGTATTTATGAAACATTTCTTTAGAAAGAAA 960
Qy      961 ACAAATTTCTTGGAGAAACAGGTTTCAGAGTATCAGCGTTTTCAGGTTTTCAGGAAACC 1020
Db      961 ACAAATTTCTTGGAGAAACAGGTTTCAGAGTATCAGCGTTTTCAGGTTTTCAGGAAACC 1020
Qy      1021 ACAGATAACGCTCTTCACTTGGATGAGATTTT 1053
Db      1021 ACAGATAACGCTCTTCACTTGGATGAGATTTT 1053

RESULT 12
PCT-US05-09639-22
; Sequence 22, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US05-09639-22

Query Match          99.4%; Score 1046.6; DB 1; Length 1056;
Best Local Similarity 99.6%; Pred. No. 6.3e-269;
Matches 1049; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy      1 ATGGGGCAGCCGGAAGGCGGAGGCGGCGGCTGGATCAGGATCAGAGATCATCTTCA 60
Db      1 ATGGGGCAGCCGGAAGGCGGAGGCGGCGGCTGGATCAGGATCAGAGATCATCTTCA 60
Qy      61 GACACCAACGAAAGTGAATTAAGTCAATGAAGCCACTCTCTAAGAAAGTCTCTCG 120
Db      61 GACACCAACGAAAGTGAATTAAGTCAATGAAGCCACTCTCTAAGAAAGTCTCTCG 120
Qy      121 CGGTTTGTCTATCTTCCAAATCCAGTACCTCTGATTTTGGAAAATGTATAAACAGGCACAG 180
Db      121 CGGTTTGTCTATCTTCCAAATCCAGTACCTCTGATTTTGGAAAATGTATAAACAGGCACAG 180
Qy      181 GCTTCTCTTGGACAGAGAGGTTGTATCAAGATCTCCCTCACTGGAACAAG 240
Db      181 GCTTCTCTTGGACAGAGAGGTTGTATCAAGATCTCCCTCACTGGAACAAG 240
Qy      241 CTTAAAGCAGATGAGAAGTACTTCATCTCTCATCTTACCTTTTTCGAGCCAGTGAT 300
Db      241 CTTAAAGCAGATGAGAAGTACTTCATCTCTCATCTTACCTTTTTCGAGCCAGTGAT 300
Qy      301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 360
Db      301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 360
Qy      361 CGCTGTTTCTATGGCTTTTCAATTTCTCATCGAATTTTCACTCAGAGATGTACAGTTTG 420
Db      361 CGCTGTTTCTATGGCTTTTCAATTTCTCATCGAATTTTCACTCAGAGATGTATTTAGTTG 420
Qy      421 CTGATAGACATTAATCAGAGATCCCAAGAAAGGGAATTTTATTTATGCAATTGAA 480
Db      421 CTGATAGACATTAATCAGAGATCCCAAGAAAGGGAATTTTATTTATGCAATTGAA 480
Qy      481 ACCATGCCCTATGTTAAGAAAAGAGAGATTTGGGCTTGGATGGATAGCAGATAGAAA 540
Db      481 ACCATGCCCTATGTTAAGAAAAGAGAGATTTGGGCTTGGATGGATAGCAGATAGAAA 540
```

```
Qy      541 TCTACTTTTTCGGGAAAGAGTGGTGGCCTTTGCTGCTGTAGAAAGAGTTTCTTCTCAGGA 600
Db      541 TCTACTTTTTCGGGAAAGAGTGGTGGCCTTTGCTGCTGTAGAAAGAGTTTCTTCTCAGGA 600
Qy      601 TCTTTTGTCTCTATATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
Db      601 TCTTTTGTCTCTATATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
Qy      661 AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACATTTGCTTGTGCTGATGTTCAA 720
Db      661 AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACATTTGCTTGTGCTGATGTTCAA 720
Qy      721 TACTTAGTAATAAGCCTTCAGAAAGAGGCTCAGGAGAGATCATTTGTGATGCTGTCAA 780
Db      721 TACTTAGTAATAAGCCTTCAGAAAGAGGCTCAGGAGAGATCATTTGTGATGCTGTCAA 780
Qy      781 ATTGAGCAGGAGTTTAAACAGAGCCTTCCAGTTCGCCCTCATTTGGAATGAATTCAT 840
Db      781 ATTGAGCAGGAGTTTAAACAGAGCCTTCCAGTTCGCCCTCATTTGGAATGAATTCAT 840
Qy      841 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTAATTTGGAACTTGGATTTCA 900
Db      841 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTAATTTGGAACTTGGATTTCA 900
Qy      901 AAGGTTTTTTCAGGCAGAGAAATCCTTTTGTATTTATGAAACATTTCTTTAGAAAGAAA 960
Db      901 AAGGTTTTTTCAGGCAGAGAAATCCTTTTGTATTTATGAAACATTTCTTTAGAAAGAAA 960
Qy      961 ACAAATTTCTTGGAGAAACAGGTTTCAGAGTATCAGCGTTTTCAGGTTTTCAGGAAACC 1020
Db      961 ACAAATTTCTTGGAGAAACAGGTTTCAGAGTATCAGCGTTTTCAGGTTTTCAGGAAACC 1020
Qy      1021 ACAGATAACGCTCTTCACTTGGATGAGATTTT 1053
Db      1021 ACAGATAACGCTCTTCACTTGGATGAGATTTT 1053

RESULT 13
PCT-US05-09639-24
; Sequence 24, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US05-09639-24

Query Match          99.4%; Score 1046.6; DB 1; Length 1056;
Best Local Similarity 99.6%; Pred. No. 6.3e-269;
Matches 1049; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy      1 ATGGGGCAGCCGGAAGGCGGAGGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db      1 ATGGGGCAGCCGGAAGGCGGAGGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Qy      61 GACACCAACGAAAGTGAATTAAGTCAATGAAGCCACTCTCTAAGAAAGTCTCTCG 120
Db      61 GACACCAACGAAAGTGAATTAAGTCAATGAAGCCACTCTCTAAGAAAGTCTCTCG 120
Qy      121 CGGTTTGTCTATCTTCCAAATCCAGTACCTCTGATTTTGGAAAATGTATAAACAGGCACAG 180
Db      121 CGGTTTGTCTATCTTCCAAATCCAGTACCTCTGATTTTGGAAAATGTATAAACAGGCACAG 180
Qy      181 GCTTCTCTTGGACAGAGAGGTTGTATCAAGATCTCCCTCACTGGAACAAG 240
Db      181 GCTTCTCTTGGACAGAGAGGTTGTATCAAGATCTCCCTCACTGGAACAAG 240
Qy      241 CTTAAAGCAGATGAGAAGTACTTCATCTCTCATCTTACCTTTTTCGAGCCAGTGAT 300
Db      241 CTTAAAGCAGATGAGAAGTACTTCATCTCTCATCTTACCTTTTTCGAGCCAGTGAT 300
Qy      301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 360
Db      301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 360
Qy      361 CGCTGTTTCTATGGCTTTTCAATTTCTCATCGAATTTTCACTCAGAGATGTACAGTTTG 420
Db      361 CGCTGTTTCTATGGCTTTTCAATTTCTCATCGAATTTTCACTCAGAGATGTATTTAGTTG 420
Qy      421 CTGATAGACATTAATCAGAGATCCCAAGAAAGGGAATTTTATTTATGCAATTGAA 480
Db      421 CTGATAGACATTAATCAGAGATCCCAAGAAAGGGAATTTTATTTATGCAATTGAA 480
Qy      481 ACCATGCCCTATGTTAAGAAAAGAGAGATTTGGGCTTGGATGGATAGCAGATAGAAA 540
Db      481 ACCATGCCCTATGTTAAGAAAAGAGAGATTTGGGCTTGGATGGATAGCAGATAGAAA 540
```

```
181 GCTTCCTCTGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAAACAAG 240
Db      |||
181 GCTTCCTCTGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAAACAAG 240
Qy      |||
241 CTTAAAGCAGATGAGAAGTACTTCTCTCTCACATCTTAGCCCTTTTTCGACGCAGTGAT 300
Db      |||
241 CTTAAAGCAGATGAGAAGTACTTCTCTCTCACATCTTAGCCCTTTTTCGACGCAGTGAT 300
Qy      |||
301 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTCCAGAGGCT 360
Db      |||
301 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTCCAGAGGCT 360
Qy      |||
361 CGCTGTTTCTTAGGCTTTCAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db      |||
361 CGCTGTTTCTTAGGCTTTCAAATTTCTCATCGAGAATGTTCACTCAGAGATGTGAGTTTG 420
Qy      |||
421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAAATGCAATTTGAA 480
Db      |||
421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAAATGCAATTTGAA 480
Qy      |||
481 ACCATGCCCTATGTTAAGAAAGAGATTTGGCGCTTCGATGGATAGCAGATAGAAA 540
Db      |||
481 ACCATGCCCTATGTTAAGAAAGAGATTTGGCGCTTCGATGGATAGCAGATAGAAA 540
Qy      |||
541 TCTACTTTTGGGAAAGAGTGGTGGCGCTTTGCTGTAGAGAGATTTTCTCTCAGGA 600
Db      |||
541 TCTACTTTTGGGAAAGAGTGGTGGCGCTTTGCTGTAGAGAGATTTTCTCTCAGGA 600
Qy      |||
601 TCTTTTGTGCTATATTTCTGGCTTAAAGAAAGAGAGTCTTATGCGAGACTCACTTTTTC 660
Db      |||
601 TCTTTTGTGCTATATTTCTGGCTTAAAGAAAGAGAGTCTTATGCGAGACTCACTTTTTC 660
Qy      |||
661 AATGAACCTCATCAGCAGATGAGAGACTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db      |||
661 AATGAACCTCATCAGCAGATGAGAGACTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Qy      |||
721 TACTTAGTAAATAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGTGATGCTGTCAA 780
Db      |||
721 TACTTAGTAAATAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGTGATGCTGTCAA 780
Qy      |||
781 ATTGACGAGGTTTTTAAACAGAGCCTTGCCAGTTGGCCTCATTCGAAATGCAATTT 840
Db      |||
781 ATTGACGAGGTTTTTAAACAGAGCCTTGCCAGTTGGCCTCATTCGAAATGCAATTT 840
Qy      |||
841 TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTAATTGGAACCTTGGAACTCTCA 900
Db      |||
841 TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTAATTGGAACCTTGGAACTCTCA 900
Qy      |||
901 AAGGTTTTTCAGCAGAAATCCCTTTGATTTTATGGAACCAATTTCTTTAGAGGAAA 960
Db      |||
901 AAGGTTTTTCAGCAGAAATCCCTTTGATTTTATGGAACCAATTTCTTTAGAGGAAA 960
Qy      |||
961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTTGCAGTTATGGCAGAAACC 1020
Db      |||
961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTTGCAGTTATGGCAGAAACC 1020
Qy      |||
1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053
Db      |||
1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053
```

RESULT 14

```
US-11-088-686-22
; Sequence 22, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; PRIOR FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-11-088-686-22
```

Query Match

```
99.4%; Score 1046.6; DB 27; Length 1056;
Best Local Similarity 99.6%; Pred. No. 6.3e-269;
Matches 1049; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
Qy      1 ATGGCGACCCGGAAGCGCGCGGCTGGATCAGGATGAGAGATCACTTTCA 60
Db      |||
Qy      1 ATGGCGACCCGGAAGCGCGCGGCTGGATCAGGATGAGAGATCACTTTCA 60
Db      |||
Qy      61 GACACCAAGAAAGTGAATTAAGTCAATTAAGAGCCACTCTTAAGAAAGATTTCTGC 120
Db      |||
Qy      61 GACACCAAGAAAGTGAATTAAGTCAATTAAGAGCCACTCTTAAGAAAGATTTCTGC 120
Db      |||
Qy      121 CGGTTTGTCTATCTTTCCAAATCCAGTACCTGATATTTGGAAAATGTATAAACAGGCACAG 180
Db      |||
Qy      121 CGGTTTGTCTATCTTTCCAAATCCAGTACCTGATATTTGGAAAATGTATAAACAGGCACAG 180
Db      |||
Qy      181 GCTTCTCTTCTGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAAACAAG 240
Db      |||
Qy      181 GCTTCTCTTCTGACAGCAGAGAGGTTGACTTATCAAGAGATCTCCCTCACTGGAAACAAG 240
Db      |||
Qy      241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACATCTTAGCCCTTTTTCGACGCAGTGAT 300
Db      |||
Qy      241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACATCTTAGCCCTTTTTCGACGCAGTGAT 300
Db      |||
Qy      301 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTCCAGAGGCT 360
Db      |||
Qy      301 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTCCAGAGGCT 360
Db      |||
Qy      361 CGCTGTTTCTTAGGCTTTCAAATTTCTCATCGAGAATGTTTCACTCAGAGATGTACAGTTTG 420
Db      |||
Qy      361 CGCTGTTTCTTAGGCTTTCAAATTTCTCATCGAGAATGTTTCACTCAGAGATGTAGTTTG 420
Db      |||
Qy      421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAAATGCAATTTGAA 480
Db      |||
Qy      421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAAATGCAATTTGAA 480
Db      |||
Qy      481 ACCATGCCCTATGTTAAGAAAGAGATTTGGCGCTTCGATGGATAGCAGATAGAAA 540
Db      |||
Qy      481 ACCATGCCCTATGTTAAGAAAGAGATTTGGCGCTTCGATGGATAGCAGATAGAAA 540
Db      |||
Qy      541 TCTACTTTTGGGAAAGAGTGGTGGCGCTTTGCTGTAGAGAGATTTTCTTCTCAGGA 600
Db      |||
Qy      541 TCTACTTTTGGGAAAGAGTGGTGGCGCTTTGCTGTAGAGAGATTTTCTTCTCAGGA 600
Db      |||
Qy      601 TCTTTTGTGCTATATTTCTGGCTTAAAGAAAGAGGCTTATGCGAGACTCACTTTTTC 660
Db      |||
Qy      601 TCTTTTGTGCTATATTTCTGGCTTAAAGAAAGAGGCTTATGCGAGACTCACTTTTTC 660
Db      |||
Qy      661 AATGAACCTCATCAGCAGATGAGAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db      |||
Qy      661 AATGAACCTCATCAGCAGATGAGAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db      |||
Qy      721 TACTTAGTAAATAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGTGATGCTGTCAA 780
Db      |||
Qy      721 TACTTAGTAAATAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGTGATGCTGTCAA 780
Db      |||
Qy      781 ATTGACGAGGTTTTTAAACAGAGCCTTGCCAGTTGGCCTCATTCGAAATGCAATTT 840
Db      |||
Qy      781 ATTGACGAGGTTTTTAAACAGAGCCTTGCCAGTTGGCCTCATTCGAAATGCAATTT 840
Db      |||
Qy      841 TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTAATTGGAACCTTGGAACTCTCA 900
Db      |||
Qy      841 TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTAATTGGAACCTTGGAACTCTCA 900
Db      |||
Qy      901 AAGGTTTTTCAGCAGAAATCCCTTTGATTTTATGGAACCAATTTCTTTAGAGGAAA 960
Db      |||
Qy      901 AAGGTTTTTCAGCAGAAATCCCTTTGATTTTATGGAACCAATTTCTTTAGAGGAAA 960
Db      |||
```


Db 901 AAGGTTTTTCAGCAGAAAATCCTTTTGATTTTATGGAATAAATTTCTTTTAGAAGAAAA 960
QY 961 ACAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC 1020
Db 961 ACAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC 1020
QY 1021 ACAGATAACGTTCTTACCTTTGGATGAGATTTT 1053
Db 1021 ACAGATAACGTTCTTACCTTTGGATGAGATTTT 1053

RESULT 15
US-11-088-686-24
; Sequence 24, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yip, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-11-088-686-24

Query Match 99.4%; Score 1046.6; DB 27; Length 1056;
Best Local Similarity 99.6%; Pred. No. 6.3e-269;
Matches 1049; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAGCGCGGCTGGATCAGGATCAGAGATCATCTTCA 60
Db 1 ATGGGCGACCCGGAAGCGCGGCTGGATCAGGATCAGAGATCATCTTCA 60

QY 61 GACACCAAGAAAGTGAATTAAGTCAATGAAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAAGAAAGTGAATTAAGTCAATGAAGCCACTCTTAAGAAAGAGTTCTCGC 120

QY 121 CGTTTGTCTATCTTTCCATCCAGTACCCTGATATTTGGAAATGATTAACAGGACAG 180
Db 121 CGTTTGTCTATCTTTCCATCCAGTACCCTGATATTTGGAAATGATTAACAGGACAG 180

QY 181 GCTTCCTCTTGGACAGCAGAGAGTTGACTTATCAAGGATCTCCCTCACTGGAACAAG 240
Db 181 GCTTCCTCTTGGACAGCAGAGAGTTGACTTATCAAGGATCTCCCTCACTGGAACAAG 240

QY 241 CTTAAGCAGATGAGAAAGTACTTCTCTACATCTTACCTTTTTCAGCCAGTGTAT 300
Db 241 CTTAAGCAGATGAGAAAGTACTTCTCTACATCTTACCTTTTTCAGCCAGTGTAT 300

QY 301 GGAATTTGAATGAATAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCT 360
Db 301 GGAATTTGAATGAATAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGCT 360

QY 361 CGCTGTTTCTATGGCTTTTCAAAATTCATCGAGATGTTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGGCTTTTCAAAATTCATCGAGATGTTTCACTCAGAGATGTACAGTTTG 420

QY 421 CTGATAGACACTTATACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTGAA 480
Db 421 CTGATAGACACTTATACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTGAA 480

QY 481 ACCATGCCCTATGTTAAGAAAGAGATTTGGCCCTTGGATGGATAGCAGATAGAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAGAGATTTGGCCCTTGGATGGATAGCAGATAGAAA 540

QY 541 TCTACTTTTGGGAAAGAGTGTGGCTTTTGTGCTGTAGAGAGAGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGTGGCTTTTGTGCTGTAGAGAGAGTTTCTTCTCAGGA 600

Search completed: October 30, 2005, 06:01:28
Job time : 1375.5 secs

QY 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACATTTTTC 660
Db 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACATTTTTC 660

QY 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCTGATGTTCCAA 720
Db 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCTGATGTTCCAA 720

QY 721 TACTTAGTAATTAAGCCTTCAGAGAAAGGCTCAGGAGAGATCATTTGTTGATGCTGTCAA 780
Db 721 TACTTAGTAATTAAGCCTTCAGAGAAAGGCTCAGGAGAGATCATTTGTTGATGCTGTCAA 780

QY 781 ATTGAGCAGGAGTTTAAACAGAGAGCTTGGCAGTTGGCCCTCATTTGGAATGAATTTGCATT 840
Db 781 ATTGAGCAGGAGTTTAAACAGAGAGCTTGGCAGTTGGCCCTCATTTGGAATGAATTTGCATT 840

QY 841 TTGATGAACACAGTACATTGAGTTTGTAGCTGACAGATTAATTGTGGAACTTTGGATTTCTCA 900
Db 841 TTGATGAACACAGTACATTGAGTTTGTAGCTGACAGATTAATTGTGGAACTTTGGATTTCTCA 900

QY 901 AAGGTTTTTTCAGGCGAGAAAATCCTTTTGAATTTTATGGAACAATTTCTTTAGAAGAAAA 960
Db 901 AAGGTTTTTTCAGGCGAGAAAATCCTTTTGAATTTTATGGAACAATTTCTTTAGAAGAAAA 960

QY 961 ACAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTATGSCAGAAACC 1020
Db 961 ACAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTATGSCAGAAACC 1020

QY 1021 ACAGATAACGTTCTTTCACCTTTGGATGAGATTTT 1053
Db 1021 ACAGATAACGTTCTTTCACCTTTGGATGAGATTTT 1053

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2005, 06:55:26 ; Search time 177 Seconds
(without alignments)
2316.216 Million cell updates/sec

Title: US-10-698-228-1

Perfect score: 1821

Sequence: 1 MGDPRPEAAGLDQDERSSS.....QRFVMAETTDNVFTLDADF 351

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 6959266 seqs, 1168006243 residues

Total number of hits satisfying chosen parameters: 6959266

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Pending Patents_AA_Main:*
- 1: /cgn2_6/ptodata/1/paa/US06 COMB.pap.*
 - 2: /cgn2_6/ptodata/1/paa/US06 COMB.pap.*
 - 3: /cgn2_6/ptodata/1/paa/US07 COMB.pap.*
 - 4: /cgn2_6/ptodata/1/paa/US08 COMB.pap.*
 - 5: /cgn2_6/ptodata/1/paa/US08 COMB.pap.*
 - 6: /cgn2_6/ptodata/1/paa/US08 COMB.pap.*
 - 7: /cgn2_6/ptodata/1/paa/US08 COMB.pap.*
 - 8: /cgn2_6/ptodata/1/paa/US08 COMB.pap.*
 - 9: /cgn2_6/ptodata/1/paa/US08 COMB.pap.*
 - 10: /cgn2_6/ptodata/1/paa/US08 COMB.pap.*
 - 11: /cgn2_6/ptodata/1/paa/US08 COMB.pap.*
 - 12: /cgn2_6/ptodata/1/paa/US08 COMB.pap.*
 - 13: /cgn2_6/ptodata/1/paa/US08 COMB.pap.*
 - 14: /cgn2_6/ptodata/1/paa/US08 COMB.pap.*
 - 15: /cgn2_6/ptodata/1/paa/US08 COMB.pap.*
 - 16: /cgn2_6/ptodata/1/paa/US08 COMB.pap.*
 - 17: /cgn2_6/ptodata/1/paa/US08 COMB.pap.*
 - 18: /cgn2_6/ptodata/1/paa/US08 COMB.pap.*
 - 19: /cgn2_6/ptodata/1/paa/US08 COMB.pap.*
 - 20: /cgn2_6/ptodata/1/paa/US08 COMB.pap.*
 - 21: /cgn2_6/ptodata/1/paa/US08 COMB.pap.*
 - 22: /cgn2_6/ptodata/1/paa/US08 COMB.pap.*
 - 23: /cgn2_6/ptodata/1/paa/US08 COMB.pap.*
 - 24: /cgn2_6/ptodata/1/paa/US08 COMB.pap.*
 - 25: /cgn2_6/ptodata/1/paa/US08 COMB.pap.*
 - 26: /cgn2_6/ptodata/1/paa/US08 COMB.pap.*
 - 27: /cgn2_6/ptodata/1/paa/US08 COMB.pap.*
 - 28: /cgn2_6/ptodata/1/paa/US08 COMB.pap.*
 - 29: /cgn2_6/ptodata/1/paa/US08 COMB.pap.*
 - 30: /cgn2_6/ptodata/1/paa/US08 COMB.pap.*
 - 31: /cgn2_6/ptodata/1/paa/US08 COMB.pap.*
 - 32: /cgn2_6/ptodata/1/paa/US08 COMB.pap.*
 - 33: /cgn2_6/ptodata/1/paa/US08 COMB.pap.*
 - 34: /cgn2_6/ptodata/1/paa/US08 COMB.pap.*
 - 35: /cgn2_6/ptodata/1/paa/US08 COMB.pap.*
 - 36: /cgn2_6/ptodata/1/paa/US08 COMB.pap.*
 - 37: /cgn2_6/ptodata/1/paa/US08 COMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1821	100.0	351	22	US-09-787-491B-11	Sequence 11, Appl
2	1821	100.0	351	22	US-09-791-537-77474	Sequence 77474, A
3	1821	100.0	351	27	US-10-170-205E-19045	Sequence 19045, A
4	1821	100.0	351	32	US-10-698-228-1	Sequence 1, Appl
5	1821	100.0	351	35	US-10-990-328-7547	Sequence 7547, Ap
6	1821	100.0	351	37	US-60-128-660-7	Sequence 7, Appl
7	1821	100.0	351	37	US-60-449-629-94	Sequence 94, Appl
8	1821	100.0	351	37	US-60-505-218-322	Sequence 322, App
9	1821	100.0	366	22	US-09-791-537-27671	Sequence 27671, A
10	1817	99.8	351	20	US-09-629-469A-12811	Sequence 12811, A
11	1817	99.8	351	22	US-09-791-537-112185	Sequence 112185, A
12	1817	99.8	351	35	US-10-917-503-12811	Sequence 12811, A
13	1454	79.8	389	1	PCT-US02-10824-143	Sequence 143, App
14	1454	79.8	389	1	PCT-US03-17409-171	Sequence 171, App
15	1454	79.8	389	10	US-08-663-617-33	Sequence 33, Appl
16	1454	79.8	389	12	US-08-837-599-33	Sequence 33, Appl
17	1454	79.8	389	22	US-09-791-537-88179	Sequence 88179, A
18	1454	79.8	389	27	US-10-170-205E-8018	Sequence 8018, Ap
19	1454	79.8	389	30	US-10-408-765-559	Sequence 559, App
20	1454	79.8	389	30	US-10-408-765A-559	Sequence 559, App
21	1454	79.8	389	32	US-10-698-228-5	Sequence 5, Appl
22	1454	79.8	389	33	US-10-733-878-457	Sequence 457, App
23	1454	79.8	389	33	US-10-756-149-4899	Sequence 4899, Ap
24	1454	79.8	389	35	US-10-940-774-6016	Sequence 6016, Ap
25	1454	79.8	389	37	US-60-389-987-559	Sequence 559, App
26	1454	79.8	389	37	US-60-412-418-559	Sequence 559, App
27	1454	79.8	389	37	US-60-452-680-20616	Sequence 20616, A
28	1454	79.8	389	37	US-60-568-073-984	Sequence 984, App
29	1454	79.8	413	1	PCT-US00-05882-1347	Sequence 1347, Ap
30	1454	79.8	413	24	US-09-525-301-1347	Sequence 1347, Ap
31	1454	79.8	449	35	US-10-990-328-12349	Sequence 12349, A
32	1454	79.8	449	35	US-10-990-328-12350	Sequence 12350, A
33	1454	79.8	453	1	PCT-US01-04926A-686	Sequence 686, App
34	1454	79.8	453	1	PCT-US01-08631-45646	Sequence 45646, A
35	1454	79.8	453	28	US-10-220-335-686	Sequence 686, App
36	1454	79.8	453	28	US-10-290-752-686	Sequence 686, App
37	1454	79.8	453	35	US-10-940-774-7896	Sequence 7896, Ap
38	1443	79.2	390	10	US-08-663-617-35	Sequence 35, Appl
39	1443	79.2	390	12	US-08-837-599-35	Sequence 35, Appl
40	1443	79.2	390	22	US-09-791-537-11257	Sequence 11257, A
41	1425.5	78.3	386	22	US-09-791-537-114288	Sequence 114288, A
42	1425.5	78.3	386	30	US-10-403-571-76	Sequence 76, Appl
43	1425.5	78.3	386	30	US-10-403-571-76	Sequence 76, Appl
44	1421	78.0	386	22	US-09-791-537-36076	Sequence 36076, A
45	1349.5	74.1	399	22	US-09-791-537-1129	Sequence 1129, Ap

ALIGNMENTS

RESULT 1

- US-09-787-491B-11
- ; Sequence 11, Application US/09787491B
 - ; GENERAL INFORMATION:
 - ; APPLICANT: INCYTE GENOMICS, INC.; TANG, Y. Tom;
 - ; APPLICANT: CORLEY, Neil C.; GUEGLER, Karl J.;
 - ; APPLICANT: GORGONE, Gina A.; PATTERSON, Chandra;
 - ; APPLICANT: HILLMAN, Jennifer L.; BAUGHN, Mariah R.;
 - ; APPLICANT: LAL, Preeti; AZIMZAI, Yalda;
 - ; APPLICANT: YUE, Henry; YANG, Junming
 - ; TITLE OF INVENTION: RNA-ASSOCIATED PROTEINS
 - ; FILE REFERENCE: PF-0600 USN
 - ; CURRENT APPLICATION NUMBER: US/09787, 491B
 - ; CURRENT FILING DATE: 2002-11-04
 - ; PRIOR APPLICATION NUMBER: PCT/US99/21688
 - ; PRIOR FILING DATE: 1999-09-17
 - ; PRIOR APPLICATION NUMBER: US 60/128,660
 - ; PRIOR FILING DATE: 1999-04-08
 - ; PRIOR APPLICATION NUMBER: US 60/069,391

```
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: US 60/183,025
; PRIOR FILING DATE: 1998-09-22
; PRIOR APPLICATION NUMBER: US 60/155,246
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: US 09/158,720
; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PERL Program
; SEQ ID NO 11
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 2604449CD1
US-09-787-491B-11

Query Match      100.0%; Score 1821; DB 22; Length 351;
Best Local Similarity 100.0%; Pred. No. 1.3e-180;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPDIWKMYKQAA 60
Db 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPDIWKMYKQAA 60
QY 61 ASFTAAEVDLSKDLPHWNKLKADEKYFISHILAFPAASDGI VNNELVERFSQEVQVPEA 120
Db 61 ASFTAAEVDLSKDLPHWNKLKADEKYFISHILAFPAASDGI VNNELVERFSQEVQVPEA 120
QY 121 RCFYGFQIILIENHSEMYSLIDTYIRDPKKRFFLNAIETMPYVKKKADWALRWIADRK 180
Db 121 RCFYGFQIILIENHSEMYSLIDTYIRDPKKRFFLNAIETMPYVKKKADWALRWIADRK 180
QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGKGLMPLGTF SNELISRDEGLHCDFACLMFQ 240
Db 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGKGLMPLGTF SNELISRDEGLHCDFACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLVGLGFS 300
Db 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLVGLGFS 300
QY 301 KVFQAEPPDFMENISLEGKTNFEKRVSEYQRFVAVMAETTDNVFTLDADF 351
Db 301 KVFQAEPPDFMENISLEGKTNFEKRVSEYQRFVAVMAETTDNVFTLDADF 351

RESULT 2
US-09-791-537-77474
; Sequence 77474, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 77474
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-77474

Query Match      100.0%; Score 1821; DB 22; Length 351;
Best Local Similarity 100.0%; Pred. No. 1.3e-180;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPDIWKMYKQAA 60
Db 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPDIWKMYKQAA 60
QY 61 ASFTAAEVDLSKDLPHWNKLKADEKYFISHILAFPAASDGI VNNELVERFSQEVQVPEA 120
Db 61 ASFTAAEVDLSKDLPHWNKLKADEKYFISHILAFPAASDGI VNNELVERFSQEVQVPEA 120
QY 121 RCFYGFQIILIENHSEMYSLIDTYIRDPKKRFFLNAIETMPYVKKKADWALRWIADRK 180
Db 121 RCFYGFQIILIENHSEMYSLIDTYIRDPKKRFFLNAIETMPYVKKKADWALRWIADRK 180
QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGKGLMPLGTF SNELISRDEGLHCDFACLMFQ 240
Db 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGKGLMPLGTF SNELISRDEGLHCDFACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLVGLGFS 300
Db 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLVGLGFS 300
QY 301 KVFQAEPPDFMENISLEGKTNFEKRVSEYQRFVAVMAETTDNVFTLDADF 351
Db 301 KVFQAEPPDFMENISLEGKTNFEKRVSEYQRFVAVMAETTDNVFTLDADF 351

RESULT 3
US-10-170-205E-19045
; Sequence 19045, Application US/10170205E
; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
; FILE REFERENCE: CL001381
; CURRENT APPLICATION NUMBER: US/10/170,205E
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19045
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-19045

Query Match      100.0%; Score 1821; DB 27; Length 351;
Best Local Similarity 100.0%; Pred. No. 1.3e-180;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPDIWKMYKQAA 60
Db 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPDIWKMYKQAA 60
QY 61 ASFTAAEVDLSKDLPHWNKLKADEKYFISHILAFPAASDGI VNNELVERFSQEVQVPEA 120
Db 61 ASFTAAEVDLSKDLPHWNKLKADEKYFISHILAFPAASDGI VNNELVERFSQEVQVPEA 120
QY 121 RCFYGFQIILIENHSEMYSLIDTYIRDPKKRFFLNAIETMPYVKKKADWALRWIADRK 180
Db 121 RCFYGFQIILIENHSEMYSLIDTYIRDPKKRFFLNAIETMPYVKKKADWALRWIADRK 180
QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGKGLMPLGTF SNELISRDEGLHCDFACLMFQ 240
Db 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGKGLMPLGTF SNELISRDEGLHCDFACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLVGLGFS 300
Db 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLVGLGFS 300
QY 301 KVFQAEPPDFMENISLEGKTNFEKRVSEYQRFVAVMAETTDNVFTLDADF 351
Db 301 KVFQAEPPDFMENISLEGKTNFEKRVSEYQRFVAVMAETTDNVFTLDADF 351

RESULT 4
US-10-698-228-1
; Sequence 1, Application US/10698228
```

GENERAL INFORMATION:
APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
TITLE OF INVENTION: New Protein and its DNA
FILE REFERENCE: 2619WOOP
CURRENT APPLICATION NUMBER: US/10/698,228
CURRENT FILING DATE: 2003-10-30
PRIOR APPLICATION NUMBER: US/10/019,733
PRIOR FILING DATE: 2001-12-28
PRIOR APPLICATION NUMBER: JP 11-181131
PRIOR FILING DATE: 1999-06-28
PRIOR APPLICATION NUMBER: JP 11-192391
PRIOR FILING DATE: 1999-07-06
PRIOR APPLICATION NUMBER: JP 2000-017770
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 14
SEQ ID NO 1
LENGTH: 351
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION:
US-10-698-228-1

Query Match 100.0%; Score 1821; DB 32; Length 351;
Best Local Similarity 100.0%; Pred. No. 1.3e-180;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVFPPIQYDPDIWKMYKQAO 60
DB 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVFPPIQYDPDIWKMYKQAO 60
QY 61 ASFWTAEEVDLSKDLPHWNKLKADKEYFISHILAFFAASDGIWNLVERFSQEVQVPEA 120
DB 61 ASFWTAEEVDLSKDLPHWNKLKADKEYFISHILAFFAASDGIWNLVERFSQEVQVPEA 120
QY 121 RCFYGFQILLIENVHSEMYSLIDTYIRDPKREFFLNAETMPYVKKADWALRWIADRK 180
DB 121 RCFYGFQILLIENVHSEMYSLIDTYIRDPKREFFLNAETMPYVKKADWALRWIADRK 180
QY 181 STFGERVAFAAVEGVFFSGSFAAIFWLKRGKGLMPLGTFNSNELSRDEGLHCDFACLMFQ 240
DB 181 STFGERVAFAAVEGVFFSGSFAAIFWLKRGKGLMPLGTFNSNELSRDEGLHCDFACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLVVELGFS 300
DB 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLVVELGFS 300
QY 301 KVFOAENPPDFMENISLEGKTNFFKRVSEYQRFVMAETTDNVFTLDADF 351
DB 301 KVFOAENPPDFMENISLEGKTNFFKRVSEYQRFVMAETTDNVFTLDADF 351

RESULT 5
US-10-990-328-7547
Sequence 7547; Application US/10990328
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele
TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
FILE REFERENCE: 2004-11-17
CURRENT APPLICATION NUMBER: US/10/990,328
CURRENT FILING DATE: 2004-11-17
NUMBER OF SEQ ID NOS: 558824
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7547
LENGTH: 351
TYPE: PRT
ORGANISM: Homo sapiens
US-10-990-328-7547
Query Match 100.0%; Score 1821; DB 35; Length 351;
Best Local Similarity 100.0%; Pred. No. 1.3e-180;

Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVFPPIQYDPDIWKMYKQAO 60
DB 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVFPPIQYDPDIWKMYKQAO 60
QY 61 ASFWTAEEVDLSKDLPHWNKLKADKEYFISHILAFFAASDGIWNLVERFSQEVQVPEA 120
DB 61 ASFWTAEEVDLSKDLPHWNKLKADKEYFISHILAFFAASDGIWNLVERFSQEVQVPEA 120
QY 121 RCFYGFQILLIENVHSEMYSLIDTYIRDPKREFFLNAETMPYVKKADWALRWIADRK 180
DB 121 RCFYGFQILLIENVHSEMYSLIDTYIRDPKREFFLNAETMPYVKKADWALRWIADRK 180
QY 181 STFGERVAFAAVEGVFFSGSFAAIFWLKRGKGLMPLGTFNSNELSRDEGLHCDFACLMFQ 240
DB 181 STFGERVAFAAVEGVFFSGSFAAIFWLKRGKGLMPLGTFNSNELSRDEGLHCDFACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLVVELGFS 300
DB 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLVVELGFS 300
QY 301 KVFOAENPPDFMENISLEGKTNFFKRVSEYQRFVMAETTDNVFTLDADF 351
DB 301 KVFOAENPPDFMENISLEGKTNFFKRVSEYQRFVMAETTDNVFTLDADF 351
RESULT 6
US-60-128-660-7
Sequence 7; Application US/60128660
GENERAL INFORMATION:
APPLICANT: Lal, Preeti
APPLICANT: Tang, Y. Tom
APPLICANT: Hillman, Jennifer L.
APPLICANT: Yue, Henry
APPLICANT: Yang, Junning
APPLICANT: Baughn, Mariah R.
TITLE OF INVENTION: RNA-ASSOCIATED PROTEINS
FILE REFERENCE: PF-0682 P
CURRENT APPLICATION NUMBER: US/60/128,660
CURRENT FILING DATE: 1999-04-08
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PERL Program
SEQ ID NO 7
LENGTH: 351
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: 2604449
US-60-128-660-7
Query Match 100.0%; Score 1821; DB 37; Length 351;
Best Local Similarity 100.0%; Pred. No. 1.3e-180;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVFPPIQYDPDIWKMYKQAO 60
DB 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVFPPIQYDPDIWKMYKQAO 60
QY 61 ASFWTAEEVDLSKDLPHWNKLKADKEYFISHILAFFAASDGIWNLVERFSQEVQVPEA 120
DB 61 ASFWTAEEVDLSKDLPHWNKLKADKEYFISHILAFFAASDGIWNLVERFSQEVQVPEA 120
QY 121 RCFYGFQILLIENVHSEMYSLIDTYIRDPKREFFLNAETMPYVKKADWALRWIADRK 180
DB 121 RCFYGFQILLIENVHSEMYSLIDTYIRDPKREFFLNAETMPYVKKADWALRWIADRK 180
QY 181 STFGERVAFAAVEGVFFSGSFAAIFWLKRGKGLMPLGTFNSNELSRDEGLHCDFACLMFQ 240
DB 181 STFGERVAFAAVEGVFFSGSFAAIFWLKRGKGLMPLGTFNSNELSRDEGLHCDFACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLVVELGFS 300
DB 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLVVELGFS 300

Db 241 YLVNKPSEERVREIIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLLVELGFS 300
QY 301 KVFQENPPDFMENISLEGKTNFFKRVSEYQRFVMAETTDNVFTLDADF 351
Db 301 KVFQENPPDFMENISLEGKTNFFKRVSEYQRFVMAETTDNVFTLDADF 351
RESULT 7
US-60-449-629-94
; Sequence 94, Application US/60449629
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001449
; CURRENT APPLICATION NUMBER: US/60/449, 629
; CURRENT FILING DATE: 2003-02-04
; NUMBER OF SEQ ID NOS: 2712
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 94
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-449-629-94

Query Match 100.0%; Score 1821; DB 37; Length 351;
Best Local Similarity 100.0%; Pred. No. 1.3e-180;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYPDIMWKYKQAO 60
Db 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYPDIMWKYKQAO 60
QY 61 ASFWTAEVDLSKDLPHWNKLKADEKYFISHILAPFAASDGIWNLVRRFSQEVQVPEA 120
Db 61 ASFWTAEVDLSKDLPHWNKLKADEKYFISHILAPFAASDGIWNLVRRFSQEVQVPEA 120
QY 121 RCFYGOILLIENHSEMYSLIDTYIRDPKREFLENAIETMPYVKKKADWALRIADRK 180
Db 121 RCFYGOILLIENHSEMYSLIDTYIRDPKREFLENAIETMPYVKKKADWALRIADRK 180
QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGMLPGLTFSNELISRDEGLHCDFACLMFQ 240
Db 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGMLPGLTFSNELISRDEGLHCDFACLMFQ 240
QY 241 YLVNKPSEERVREIIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLLVELGFS 300
Db 241 YLVNKPSEERVREIIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLLVELGFS 300
QY 301 KVFQENPPDFMENISLEGKTNFFKRVSEYQRFVMAETTDNVFTLDADF 351
Db 301 KVFQENPPDFMENISLEGKTNFFKRVSEYQRFVMAETTDNVFTLDADF 351

RESULT 8
US-60-505-218-322
; Sequence 322, Application US/60505218
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001482
; CURRENT APPLICATION NUMBER: US/60/505, 218
; CURRENT FILING DATE: 2003-09-24
; NUMBER OF SEQ ID NOS: 22507
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 322
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-505-218-322

Query Match 100.0%; Score 1821; DB 37; Length 351;

Best Local Similarity 100.0%; Pred. No. 1.3e-180;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYPDIMWKYKQAO 60
Db 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYPDIMWKYKQAO 60
QY 61 ASFWTAEVDLSKDLPHWNKLKADEKYFISHILAPFAASDGIWNLVRRFSQEVQVPEA 120
Db 61 ASFWTAEVDLSKDLPHWNKLKADEKYFISHILAPFAASDGIWNLVRRFSQEVQVPEA 120
QY 121 RCFYGOILLIENHSEMYSLIDTYIRDPKREFLENAIETMPYVKKKADWALRIADRK 180
Db 121 RCFYGOILLIENHSEMYSLIDTYIRDPKREFLENAIETMPYVKKKADWALRIADRK 180
QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGMLPGLTFSNELISRDEGLHCDFACLMFQ 240
Db 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGMLPGLTFSNELISRDEGLHCDFACLMFQ 240
QY 241 YLVNKPSEERVREIIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLLVELGFS 300
Db 241 YLVNKPSEERVREIIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLLVELGFS 300
QY 301 KVFQENPPDFMENISLEGKTNFFKRVSEYQRFVMAETTDNVFTLDADF 351
Db 301 KVFQENPPDFMENISLEGKTNFFKRVSEYQRFVMAETTDNVFTLDADF 351
RESULT 9
US-09-791-537-27671
; Sequence 27671, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 27671
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-27671

Query Match 100.0%; Score 1821; DB 22; Length 366;
Best Local Similarity 100.0%; Pred. No. 1.4e-180;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYPDIMWKYKQAO 60
Db 16 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYPDIMWKYKQAO 75
QY 61 ASFWTAEVDLSKDLPHWNKLKADEKYFISHILAPFAASDGIWNLVRRFSQEVQVPEA 120
Db 76 ASFWTAEVDLSKDLPHWNKLKADEKYFISHILAPFAASDGIWNLVRRFSQEVQVPEA 135
QY 121 RCFYGOILLIENHSEMYSLIDTYIRDPKREFLENAIETMPYVKKKADWALRIADRK 180
Db 136 RCFYGOILLIENHSEMYSLIDTYIRDPKREFLENAIETMPYVKKKADWALRIADRK 195
QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGMLPGLTFSNELISRDEGLHCDFACLMFQ 240
Db 196 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGMLPGLTFSNELISRDEGLHCDFACLMFQ 255
QY 241 YLVNKPSEERVREIIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLLVELGFS 300
Db 256 YLVNKPSEERVREIIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLLVELGFS 315
QY 301 KVFQENPPDFMENISLEGKTNFFKRVSEYQRFVMAETTDNVFTLDADF 351

Db 316 KVFQAEPPDFMENISLEGKTNFFKRVSEYQRFVAVMAETTDNVFTLDADF 366

RESULT 10

US-09-629-469A-12811
; Sequence 12811, Application US/09629469A
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: SAITO, KAORU
; APPLICANT: YAMAMOTO, JUNICHI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: WAKAMATSU, AI
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: OTSUKI, TETSUJI
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH CDNA AND THEIR USE
; FILE REFERENCE: 084335/0123
; CURRENT APPLICATION NUMBER: US/09/629,469A
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: JP 1999-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 1999-300253
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/183,322
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 19025
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 12811
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-629-469A-12811

Query Match 99.8%; Score 1817; DB 20; Length 351;
Best Local Similarity 99.7%; Pred. No. 3.4e-180;
Matches 350; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDODERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPDIWKMYKQAO 60
Db 1 MGDPERPEAAGLDODERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPDIWKMYKQAO 60
QY 61 ASFWTAEVDLSKDLPHWNKLKADKEYFISHILAFFAASDGI VNNELVERFSQEVQVPEA 120
Db 61 ASFWTAEVDLSKDLPHWNKLKADKEYFISHILAFFAASDGI VNNELVERFSQEVQVPEA 120
QY 121 RCFYGFQILLIENVHSEMYSLIDTYIRDPKKREFLNAETMPYVKKKADWALRWIADRK 180
Db 121 RCFYGFQILLIENVHSEMYSLIDTYIRDPKKREFLNAETMPYVKKKADWALRWIADRK 180
QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKKRGLMPGLTFSNELISRDEGLHCDPACLMFQ 240
Db 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKKRGLMPGLTFSNELISRDEGLHCDPACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIQEFLTEALPVGLIGMNCILMKQYIEFVADRLLVELGFS 300
Db 241 YLVNKPSEERVREIIVDAVKIQEFLTEALPVGLIGMNCILMKQYIEFVADRLLVELGFS 300
QY 301 KVFQAEPPDFMENISLEGKTNFFKRVSEYQRFVAVMAETTDNVFTLDADF 351
Db 301 KVFQAEPPDFMENISLEGKTNFFKRVSEYQRFVAVMAETTDNVFTLDADF 351

RESULT 11

US-09-791-537-112185
; Sequence 112185, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patent In version 3.0
; SEQ ID NO 112185
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-112185

Query Match 99.8%; Score 1817; DB 22; Length 351;
Best Local Similarity 99.7%; Pred. No. 3.4e-180;
Matches 350; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDODERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPDIWKMYKQAO 60
Db 1 MGDPERPEAAGLDODERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPDIWKMYKQAO 60
QY 61 ASFWTAEVDLSKDLPHWNKLKADKEYFISHILAFFAASDGI VNNELVERFSQEVQVPEA 120
Db 61 ASFWTAEVDLSKDLPHWNKLKADKEYFISHILAFFAASDGI VNNELVERFSQEVQVPEA 120
QY 121 RCFYGFQILLIENVHSEMYSLIDTYIRDPKKREFLNAETMPYVKKKADWALRWIADRK 180
Db 121 RCFYGFQILLIENVHSEMYSLIDTYIRDPKKREFLNAETMPYVKKKADWALRWIADRK 180
QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKKRGLMPGLTFSNELISRDEGLHCDPACLMFQ 240
Db 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKKRGLMPGLTFSNELISRDEGLHCDPACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIQEFLTEALPVGLIGMNCILMKQYIEFVADRLLVELGFS 300
Db 241 YLVNKPSEERVREIIVDAVKIQEFLTEALPVGLIGMNCILMKQYIEFVADRLLVELGFS 300
QY 301 KVFQAEPPDFMENISLEGKTNFFKRVSEYQRFVAVMAETTDNVFTLDADF 351
Db 301 KVFQAEPPDFMENISLEGKTNFFKRVSEYQRFVAVMAETTDNVFTLDADF 351

RESULT 12

US-10-917-503-12811
; Sequence 12811, Application US/10917503
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: SAITO, KAORU
; APPLICANT: YAMAMOTO, JUNICHI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: WAKAMATSU, AI
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: OTSUKI, TETSUJI
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH CDNA AND THEIR USE
; FILE REFERENCE: 084335/0123
; CURRENT APPLICATION NUMBER: US/10/917,503
; CURRENT FILING DATE: 2004-08-13
; PRIOR APPLICATION NUMBER: US/09/629,469
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: JP 1999-248036
; PRIOR FILING DATE: 1999-07-29

```
; PRIOR APPLICATION NUMBER: JP 1999-300253
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/183,322
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 19025
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12811
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-917-503-12811

Query Match      99.8%; Score 1817; DB 35; Length 351;
Best Local Similarity 99.7%; Pred. No. 3.4e-180;
Matches 350; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVFPFIQYDPDIWKMVKQAQ 60
DB 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVFPFIQYDPDIWKMVKQAQ 60

QY 61 ASFWTAEEVDLSKDLPHWNKLKADEKFIHILAFPAASDGIWNLVERFSQEVQVPEA 120
DB 61 ASFWTAEEVDLSKDLPHWNKLKADEKFIHILAFPAASDGIWNLVERFSQEVQVPEA 120

QY 121 RCFYGFQIILNENHVSMSYLLIDTYIRDPKKEFLNATETMPYVKKKADWALRWIADRX 180
DB 121 RCFYGFQIILNENHVSMSYLLIDTYIRDPKKEFLNATETMPYVKKKADWALRWIADRX 180

QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKKRGLMPLGTFNSNELISRDEGLHCDFACLMFQ 240
DB 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKKRGLMPLGTFNSNELISRDEGLHCDFACLMFQ 240

QY 241 YLVNKPSEERREIIVDAVKIEQFLTEALPVGLIGNCILMKQYIEFVADRLLVGLGFS 300
DB 241 YLVNKPSEERREIIVDAVKIEQFLTEALPVGLIGNCILMKQYIEFVADRLLVGLGFS 300

QY 301 KVFQAEPPDFMENISLEGKTNFFEKRVSEYQRFVAVMAETTDNVFTLDADF 351
DB 301 KVFQAEPPDFMENISLEGKTNFFEKRVSEYQRFVAVMAETTDNVFTLDADF 351

RESULT 13
PCT-US02-10824-143
; Sequence 143, Application PC/TUS0210824
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Prostate Cancer Expression Profiles
; FILE REFERENCE: 9U 206 PCT
; CURRENT APPLICATION NUMBER: PCT/US02/10824
; CURRENT FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: US 60/281,732
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 60/281,731
; PRIOR FILING DATE: 2001-04-06
; NUMBER OF SEQ ID NOS: 211
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 143
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-10824-143

Query Match      79.8%; Score 1454; DB 1; Length 389;
Best Local Similarity 83.8%; Pred. No. 3.5e-142;
Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;

PCT-US02-10824-143
```

```
QY 31 EEPLLRKSRRFVFPFIQYDPDIWKMVKQAASFWTAEEVDLSKDLPHWNKLKADEKFI 90
DB 69 DEPLLRENPRFVFPPIEYHDIWQMYKKAASFWTAEEVDLSKDIQHWESLKPEERYFIS 128

QY 91 HILAFPAASDGIWNLVERFSQEVQVPEARCFYGFQIILNENHVSMSYLLIDTYIRDPK 150
DB 129 HVLAFPAASDGIWNLVERFSQEVQITEARCFYGFQIAMENIHSEMYSLIDTYIKDPK 188

QY 151 KRSFLNATETMPYVKKKADWALRWIADRXSTGERVAVFAAAGVFPSSGFAAIFWLKK 210
DB 189 EREFLNATETMPYVKKKADWALRWIGDKEATYGERVAVFAAAGVFPSSGFAAIFWLKK 248

QY 211 RGLMPGLTFSNELISRDEGLHCDFACLMFQYLVNKPSEERREIIVDAVKIEQFLTEAL 270
DB 249 RGLMPGLTFSNELISRDEGLHCDFACLMFKHLVHKPSEERREIINAVRIEQEFLTEAL 308

QY 271 PVGLIGNCILMKQYIEFVADRLLVGLGFSKVFOAENPFDPMENISLEGKTNFFEKRVSE 330
DB 309 PVKLGIMNCTLMKQYIEFVADRLMLGLGFSKVFRVENPFDPMENISLEGKTNFFEKRVGE 368

QY 331 YQRFVAVMAETTDNVFTLDADF 351
DB 369 YQRMGVMSPTENSFTLDADF 389

RESULT 14
PCT-US03-17409-171
; Sequence 171, Application PC/TUS0317409
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION
; APPLICANT: RICKERT, Paula K.
; APPLICANT: KRASNOW, Randi
; TITLE OF INVENTION: DIAGNOSTIC MARKERS FOR LUNG CANCER
; FILE REFERENCE: PA-0051 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/17409
; CURRENT FILING DATE: 2003-06-02
; PRIOR APPLICATION NUMBER: US 60/386,005
; PRIOR FILING DATE: 2002-06-04
; NUMBER OF SEQ ID NOS: 296
; SOFTWARE: PERL Program
; SEQ ID NO 171
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 000899CD1
PCT-US03-17409-171

Query Match      79.8%; Score 1454; DB 1; Length 389;
Best Local Similarity 83.8%; Pred. No. 3.5e-142;
Matches 269; Conservative 33; Mismatches 19; Indels 0; Gaps 0;

QY 31 EEPLLRKSRRFVFPFIQYDPDIWKMVKQAASFWTAEEVDLSKDLPHWNKLKADEKFI 90
DB 69 DEPLLRENPRFVFPPIEYHDIWQMYKKAASFWTAEEVDLSKDIQHWESLKPEERYFIS 128

QY 91 HILAFPAASDGIWNLVERFSQEVQVPEARCFYGFQIILNENHVSMSYLLIDTYIRDPK 150
DB 129 HVLAFPAASDGIWNLVERFSQEVQITEARCFYGFQIAMENIHSEMYSLIDTYIKDPK 188

QY 151 KRSFLNATETMPYVKKKADWALRWIADRXSTGERVAVFAAAGVFPSSGFAAIFWLKK 210
DB 189 EREFLNATETMPYVKKKADWALRWIGDKEATYGERVAVFAAAGVFPSSGFAAIFWLKK 248

QY 211 RGLMPGLTFSNELISRDEGLHCDFACLMFQYLVNKPSEERREIIVDAVKIEQFLTEAL 270
DB 249 RGLMPGLTFSNELISRDEGLHCDFACLMFKHLVHKPSEERREIINAVRIEQEFLTEAL 308

QY 271 PVGLIGNCILMKQYIEFVADRLLVGLGFSKVFOAENPFDPMENISLEGKTNFFEKRVSE 330
DB 309 PVKLGIMNCTLMKQYIEFVADRLMLGLGFSKVFRVENPFDPMENISLEGKTNFFEKRVGE 368
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2005, 06:56:26 ; Search time 44 Seconds
(without alignments)
1802.648 Million cell updates/sec

Title: US-10-698-228-1

Perfect score: 1821

Sequence: 1 MGDPERPEAGLDQDERSSS.....QRFVMAETTDNVFTLDADF 351

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 972358 seqs, 225972999 residues

Total number of hits satisfying chosen parameters: 972358

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents AA New.*

- 1: /cgn2.6/prodata/1/paa/PCT_NEW_COMB.pep.*
- 2: /cgn2.6/prodata/1/paa/US05_NEW_COMB.pep.*
- 3: /cgn2.6/prodata/1/paa/US07_NEW_COMB.pep.*
- 4: /cgn2.6/prodata/1/paa/US08_NEW_COMB.pep.*
- 5: /cgn2.6/prodata/1/paa/US09_NEW_COMB.pep.*
- 6: /cgn2.6/prodata/1/paa/US10_NEW_COMB.pep.*
- 7: /cgn2.6/prodata/1/paa/US11_NEW_COMB.pep.*
- 8: /cgn2.6/prodata/1/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1821	100.0	351	1	PCT-US05-09639-15
2	1821	100.0	351	1	PCT-US05-21650-34
3	1821	100.0	351	6	US-10-990-328A-7547
4	1821	100.0	351	7	US-11-088-686-15
5	1817	99.8	351	1	PCT-US05-09639-17
6	1817	99.8	351	7	US-11-088-686-17
7	1816	99.7	351	1	PCT-US05-09639-19
8	1816	99.7	351	7	US-11-088-686-19
9	1813	99.6	351	1	PCT-US05-09639-21
10	1813	99.6	351	7	US-11-088-686-21
11	1812	99.5	351	1	PCT-US05-09639-23
12	1812	99.5	351	1	PCT-US05-09639-25
13	1812	99.5	351	7	US-11-088-686-23
14	1812	99.5	351	7	US-11-088-686-25
15	1811	99.5	351	1	PCT-US05-09639-27
16	1811	99.5	351	7	US-11-088-686-27
17	1454	79.8	389	1	PCT-US03-10870-559
18	1454	79.8	389	1	PCT-US05-09639-1
19	1454	79.8	389	1	PCT-US05-21650-33
20	1454	79.8	389	6	US-10-940-774A-6016
21	1454	79.8	389	7	US-11-088-686-1
22	1454	79.8	449	6	US-10-990-328A-12349
23	1454	79.8	449	6	US-10-990-328A-12350
24	1454	79.8	453	6	US-10-450-763-45646
25	1454	79.8	453	6	US-10-940-774A-7896

26	1450	79.6	389	1	PCT-US05-09639-3	Sequence 3, Appli
27	1450	79.6	389	7	US-11-088-686-3	Sequence 3, Appli
28	1449	79.6	389	1	PCT-US05-09639-5	Sequence 5, Appli
29	1449	79.6	389	7	US-11-088-686-5	Sequence 5, Appli
30	1446	79.4	389	1	PCT-US05-09639-7	Sequence 7, Appli
31	1446	79.4	389	7	US-11-088-686-7	Sequence 7, Appli
32	1445	79.4	389	1	PCT-US05-09639-9	Sequence 9, Appli
33	1445	79.4	389	1	PCT-US05-09639-11	Sequence 11, Appli
34	1445	79.4	389	7	US-11-088-686-9	Sequence 9, Appli
35	1445	79.4	389	7	US-11-088-686-11	Sequence 11, Appli
36	1444	79.3	389	1	PCT-US05-09639-13	Sequence 13, Appli
37	1444	79.3	389	7	US-11-088-686-13	Sequence 13, Appli
38	1371	75.3	389	8	US-60-669-241-28508	Sequence 28508, A
39	1348	74.0	393	7	US-11-097-143-19182	Sequence 19182, A
40	1274	70.0	319	1	PCT-US05-12947-102	Sequence 102, App
41	1232.5	67.7	404	6	US-10-703-032-123588	Sequence 123588,
42	1156	63.5	330	7	US-11-251-208-421	Sequence 421, App
43	1147.5	63.0	329	7	US-11-251-208-427	Sequence 427, App
44	1128.5	62.0	327	7	US/11/241	GENERAL INFORMATI
45	1125.5	61.8	345	7	US-11-251-208-519	Sequence 519, App

ALIGNMENTS

RESULT 1
PCT-US05-09639-15
; Sequence 15, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004WO1
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-09639-15

Query Match 100.0%; Score 1821; DB 1; Length 351;
Best Local Similarity 100.0%; Pred. No. 3.3e-160;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MGDPERPEAGLDQDERSSSDTNESEITKSNEEPLLRKSSRRFVFPQYPDIMWKYKQAQ	60
Db	1	MGDPERPEAGLDQDERSSSDTNESEITKSNEEPLLRKSSRRFVFPQYPDIMWKYKQAQ	60
Qy	61	ASFWTAEEVDLSKDLPHWNKLKADEKYFISHILAFFAASDGI VNNELVERFSQEVQVPEA	120
Db	61	ASFWTAEEVDLSKDLPHWNKLKADEKYFISHILAFFAASDGI VNNELVERFSQEVQVPEA	120
Qy	121	RCFYGFQILNIENHVHSEMYSLIDITYIRDPKREFLNAIETMPYVKKKADWALRIADRK	180
Db	121	RCFYGFQILNIENHVHSEMYSLIDITYIRDPKREFLNAIETMPYVKKKADWALRIADRK	180
Qy	181	STGERVVAFAAEGVEGFFSGSFAAIFWLKXGLNPGLTFSNELISRDEGLHCDFACLMFQ	240
Db	181	STGERVVAFAAEGVEGFFSGSFAAIFWLKXGLNPGLTFSNELISRDEGLHCDFACLMFQ	240
Qy	241	YLNKKSSEERVREIIVDAVKIEQEFLEALPVGLIGNCILMKOYIEFVADRLLVELGFS	300
Db	241	YLNKKSSEERVREIIVDAVKIEQEFLEALPVGLIGNCILMKOYIEFVADRLLVELGFS	300
Qy	301	KVFQAEFPDFMENISLEGKTNFFEKRVSEYQRFVAVMAETTDNVFTLDADF	351
Db	301	KVFQAEFPDFMENISLEGKTNFFEKRVSEYQRFVAVMAETTDNVFTLDADF	351


```

RESULT 2
PCT-US05-21650-34
; Sequence 34, Application PC/TUS0521650
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: MIGFS AS MODIFIERS OF THE IGF PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX05-017
; CURRENT APPLICATION NUMBER: PCT/US05/21650
; CURRENT FILING DATE: 2005-06-27
; PRIOR APPLICATION NUMBER: 60/581,689
; PRIOR FILING DATE: 2004-06-21
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 34
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-21650-34

Query Match      100.0%; Score 1821; DB 1; Length 351;
Best Local Similarity 100.0%; Pred. No. 3.3e-160;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPDIWKMYKQAA 60
Db 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPDIWKMYKQAA 60
QY 61 ASFWTAEVDLSKDLPHWNKLKADEKYFISHILAFPAASDGIIVNENLVERFSQEVQVPEA 120
Db 61 ASFWTAEVDLSKDLPHWNKLKADEKYFISHILAFPAASDGIIVNENLVERFSQEVQVPEA 120
QY 121 RCFYGFQIILIENHVSEMYSLIDTYIRDPKREFFLNAIETMPYVKKADWALRWIADRK 180
Db 121 RCFYGFQIILIENHVSEMYSLIDTYIRDPKREFFLNAIETMPYVKKADWALRWIADRK 180
QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGMLMPLGTFPSNELISRDEGLHCDPACLMFQ 240
Db 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGMLMPLGTFPSNELISRDEGLHCDPACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFTEALPVGLIGMNCILMKQYIEFVADRLLVELGFS 300
Db 241 YLVNKPSEERVREIIVDAVKIEQEFTEALPVGLIGMNCILMKQYIEFVADRLLVELGFS 300
QY 301 KVFQAEENPFDPMENISLEGKTNFFEKRVSEYQRFVAVMAETTDNVFTLDADF 351
Db 301 KVFQAEENPFDPMENISLEGKTNFFEKRVSEYQRFVAVMAETTDNVFTLDADF 351

RESULT 4
US-11-088-686-15
; Sequence 15, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-088-686-15

Query Match      100.0%; Score 1821; DB 7; Length 351;
Best Local Similarity 100.0%; Pred. No. 3.3e-160;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPDIWKMYKQAA 60
Db 1 MGDPERPEAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYDPDIWKMYKQAA 60
QY 61 ASFWTAEVDLSKDLPHWNKLKADEKYFISHILAFPAASDGIIVNENLVERFSQEVQVPEA 120
Db 61 ASFWTAEVDLSKDLPHWNKLKADEKYFISHILAFPAASDGIIVNENLVERFSQEVQVPEA 120
QY 121 RCFYGFQIILIENHVSEMYSLIDTYIRDPKREFFLNAIETMPYVKKADWALRWIADRK 180
Db 121 RCFYGFQIILIENHVSEMYSLIDTYIRDPKREFFLNAIETMPYVKKADWALRWIADRK 180
QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGMLMPLGTFPSNELISRDEGLHCDPACLMFQ 240
Db 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGMLMPLGTFPSNELISRDEGLHCDPACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFTEALPVGLIGMNCILMKQYIEFVADRLLVELGFS 300
Db 241 YLVNKPSEERVREIIVDAVKIEQEFTEALPVGLIGMNCILMKQYIEFVADRLLVELGFS 300
QY 301 KVFQAEENPFDPMENISLEGKTNFFEKRVSEYQRFVAVMAETTDNVFTLDADF 351
Db 301 KVFQAEENPFDPMENISLEGKTNFFEKRVSEYQRFVAVMAETTDNVFTLDADF 351

RESULT 5
PCT-US05-09639-17

```

```
; Sequence 17, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; PRIOR FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-09639-17

Query Match      99.8%; Score 1817; DB 1; Length 351;
Best Local Similarity 99.7%; Pred. No. 7.9e-160;
Matches 350; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDQDSSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYPDIMWKYKQAQ 60
DB 1 MGDPERPEAAGLDQDSSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYPDIMWKYKQAQ 60
QY 61 ASFWTAEVDLSKDLPHWNKLKADEKYFISHILAPFAASDGI VNNLVERFSQEVQVPEA 120
DB 61 ASFWTAEVDLSKDLPHWNKLKADEKYFISHILAPFAASDGI VNNLVERFSQEVQVPEA 120
QY 121 RCFYGFQILLIENVHSEMYSLIDTYIRDPKPKREFLNAIETMPYVKKKADWALRWIADRK 180
DB 121 RCFYGFQILLIENVHSEMYSLIDTYIRDPKPKREFLNAIETMPYVKKKADWALRWIADRK 180
QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKKRGLMPGLTFSNELISRDEGLHCDPACLMFQ 240
DB 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKKRGLMPGLTFSNELISRDEGLHCDPACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLVELGFS 300
DB 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLVELGFS 300
QY 301 KVFQENPFDFMENISLEGKTNPFKEKRVSEYQRFVAVMAETTDNVFTLDADF 351
DB 301 KVFQENPFDFMENISLEGKTNPFKEKRVSEYQRFVAVMAETTDNVFTLDADF 351

RESULT 7
PCT-US05-09639-19
; Sequence 19, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; PRIOR FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-09639-19

Query Match      99.7%; Score 1816; DB 1; Length 351;
Best Local Similarity 99.7%; Pred. No. 9.7e-160;
Matches 350; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDQDSSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYPDIMWKYKQAQ 60
DB 1 MGDPERPEAAGLDQDSSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYPDIMWKYKQAQ 60
QY 61 ASFWTAEVDLSKDLPHWNKLKADEKYFISHILAPFAASDGI VNNLVERFSQEVQVPEA 120
DB 61 ASFWTAEVDLSKDLPHWNKLKADEKYFISHILAPFAASDGI VNNLVERFSQEVQVPEA 120
QY 121 RCFYGFQILLIENVHSEMYSLIDTYIRDPKPKREFLNAIETMPYVKKKADWALRWIADRK 180
DB 121 RCFYGFQILLIENVHSEMYSLIDTYIRDPKPKREFLNAIETMPYVKKKADWALRWIADRK 180
QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKKRGLMPGLTFSNELISRDEGLHCDPACLMFQ 240
DB 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKKRGLMPGLTFSNELISRDEGLHCDPACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLVELGFS 300
DB 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLVELGFS 300
QY 301 KVFQENPFDFMENISLEGKTNPFKEKRVSEYQRFVAVMAETTDNVFTLDADF 351
DB 301 KVFQENPFDFMENISLEGKTNPFKEKRVSEYQRFVAVMAETTDNVFTLDADF 351

RESULT 8
US-11-088-686-19
; Sequence 19, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: US/11/088,686
; PRIOR FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-088-686-17

Query Match      99.8%; Score 1817; DB 7; Length 351;
Best Local Similarity 99.7%; Pred. No. 7.9e-160;
Matches 350; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDQDSSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYPDIMWKYKQAQ 60
DB 1 MGDPERPEAAGLDQDSSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYPDIMWKYKQAQ 60
QY 61 ASFWTAEVDLSKDLPHWNKLKADEKYFISHILAPFAASDGI VNNLVERFSQEVQVPEA 120
DB 61 ASFWTAEVDLSKDLPHWNKLKADEKYFISHILAPFAASDGI VNNLVERFSQEVQVPEA 120
QY 121 RCFYGFQILLIENVHSEMYSLIDTYIRDPKPKREFLNAIETMPYVKKKADWALRWIADRK 180
DB 121 RCFYGFQILLIENVHSEMYSLIDTYIRDPKPKREFLNAIETMPYVKKKADWALRWIADRK 180
QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKKRGLMPGLTFSNELISRDEGLHCDPACLMFQ 240
DB 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKKRGLMPGLTFSNELISRDEGLHCDPACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLVELGFS 300
DB 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLLVELGFS 300
QY 301 KVFQENPFDFMENISLEGKTNPFKEKRVSEYQRFVAVMAETTDNVFTLDADF 351
DB 301 KVFQENPFDFMENISLEGKTNPFKEKRVSEYQRFVAVMAETTDNVFTLDADF 351
```

```
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-088-686-19

Query Match      99.7%; Score 1816; DB 7; Length 351;
Best Local Similarity 99.7%; Pred. No. 9.7e-160;
Matches 350; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDODRSDSSDTNESEIKNEEPLLRKSSRRFVIFPIQYDPDIWMYKQAO 60
DB 1 MGDPERPEAAGLDODRSDSSDTNESEIKNEEPLLRKSSRRFVIFPIQYDPDIWMYKQAO 60
QY 61 ASFWTAEVDLSKDLPHWNKLKADEKYFISHILAFFAASDGI VNNELVERFSQEVQVPEA 120
DB 61 ASFWTAEVDLSKDLPHWNKLKADEKYFISHILAFFAASDGI VNNELVERFSQEVQVPEA 120
QY 121 RCFYGFQILLIENHSEMYSLIDITYIRDPKKREFLNAIETMPYVKKKADWALRWIADRK 180
DB 121 RCFYGFQILLIENHSEMYSLIDITYIRDPKKREFLNAIETMPYVKKKADWALRWIADRK 180
QY 121 RCFWGFQILLIENHSEMYSLIDITYIRDPKKREFLNAIETMPYVKKKADWALRWIADRK 180
DB 121 RCFWGFQILLIENHSEMYSLIDITYIRDPKKREFLNAIETMPYVKKKADWALRWIADRK 180
QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKKRGLMPGLTFSNELISRDEGLHCDFACLMFQ 240
DB 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKKRGLMPGLTFSNELISRDEGLHCDFACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGNCILMKQYIEFVADRLLVELGFS 300
DB 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGNCILMKQYIEFVADRLLVELGFS 300
QY 301 KVFQENPPDFMENISLEGKTNFPEKRVSEYQRFVAVMAETTDNVFTLDADF 351
DB 301 KVFQENPPDFMENISLEGKTNFPEKRVSEYQRFVAVMAETTDNVFTLDADF 351

RESULT 10
US-11-088-686-21
; Sequence 21, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-088-686-21

Query Match      99.6%; Score 1813; DB 7; Length 351;
Best Local Similarity 99.4%; Pred. No. 1.8e-159;
Matches 349; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDODRSDSSDTNESEIKNEEPLLRKSSRRFVIFPIQYDPDIWMYKQAO 60
DB 1 MGDPERPEAAGLDODRSDSSDTNESEIKNEEPLLRKSSRRFVIFPIQYDPDIWMYKQAO 60
QY 61 ASFWTAEVDLSKDLPHWNKLKADEKYFISHILAFFAASDGI VNNELVERFSQEVQVPEA 120
DB 61 ASFWTAEVDLSKDLPHWNKLKADEKYFISHILAFFAASDGI VNNELVERFSQEVQVPEA 120
QY 121 RCFYGFQILLIENHSEMYSLIDITYIRDPKKREFLNAIETMPYVKKKADWALRWIADRK 180
DB 121 RCFYGFQILLIENHSEMYSLIDITYIRDPKKREFLNAIETMPYVKKKADWALRWIADRK 180
QY 121 RCFWGFQILLIENHSEMYSLIDITYIRDPKKREFLNAIETMPYVKKKADWALRWIADRK 180
DB 121 RCFWGFQILLIENHSEMYSLIDITYIRDPKKREFLNAIETMPYVKKKADWALRWIADRK 180
QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKKRGLMPGLTFSNELISRDEGLHCDFACLMFQ 240
DB 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKKRGLMPGLTFSNELISRDEGLHCDFACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGNCILMKQYIEFVADRLLVELGFS 300
DB 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGNCILMKQYIEFVADRLLVELGFS 300
QY 301 KVFQENPPDFMENISLEGKTNFPEKRVSEYQRFVAVMAETTDNVFTLDADF 351
DB 301 KVFQENPPDFMENISLEGKTNFPEKRVSEYQRFVAVMAETTDNVFTLDADF 351

RESULT 11
PCT-US05-09639-23
; Sequence 23, Application PCT/US0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
```

```
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-09639-23

Query Match          99.5%; Score 1812; DB 1; Length 351;
Best Local Similarity 99.4%; Pred. No. 2.3e-159;
Matches 349; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYPDIMWKYKQAA 60
DB 1 MGDPERPAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYPDIMWKYKQAA 60
QY 61 ASFTAEAEVDLSKDLPHWNKLKADKDYFISHILAFPAASDGI VNNLVERFSQEVQVPEA 120
DB 61 ASFTAEAEVDLSKDLPHWNKLKADKDYFISHILAFPAASDGI VNNLVERFSQEVQVPEA 120
QY 121 RCFYGFQILLIENHVHSEMYSLIDTYIRDPKREFLFNAIETMPYVKKKADWALRWIADRK 180
DB 121 RCFYGFQILLIENHVHSEMYSLIDTYIRDPKREFLFNAIETMPYVKKKADWALRWIADRK 180
QY 121 RCFYGFQILLIENHVHSEMYSLIDTYIRDPKREFLFNAIETMPYVKKKADWALRWIADRK 180
DB 121 RCFYGFQILLIENHVHSEMYSLIDTYIRDPKREFLFNAIETMPYVKKKADWALRWIADRK 180
QY 181 STFGERVVAFPAAVEGVFFSGSFAAIFWLKRGKGLMPGLTFSNELISRDEGLHCDPACLMFQ 240
DB 181 STFGERVVAFPAAVEGVFFSGSFAAIFWLKRGKGLMPGLTFSNELISRDEGLHCDPACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLVLELGS 300
DB 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLVLELGS 300
QY 301 KVFOAENPFDPMENISLEGKTNFFKRVSEYQRFVAVMAETTDNVFTLDADF 351
DB 301 KVFOAENPFDPMENISLEGKTNFFKRVSEYQRFVAVMAETTDNVFTLDADF 351

RESULT 13
US-11-088-686-23
; Sequence 23, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-088-686-23

Query Match          99.5%; Score 1812; DB 7; Length 351;
Best Local Similarity 99.4%; Pred. No. 2.3e-159;
Matches 349; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYPDIMWKYKQAA 60
DB 1 MGDPERPAAGLDQDERSSDTNESEIKSNEEPLLRKSSRRFVIFPIQYPDIMWKYKQAA 60
QY 61 ASFTAEAEVDLSKDLPHWNKLKADKDYFISHILAFPAASDGI VNNLVERFSQEVQVPEA 120
DB 61 ASFTAEAEVDLSKDLPHWNKLKADKDYFISHILAFPAASDGI VNNLVERFSQEVQVPEA 120
QY 121 RCFYGFQILLIENHVHSEMYSLIDTYIRDPKREFLFNAIETMPYVKKKADWALRWIADRK 180
DB 121 RCFYGFQILLIENHVHSEMYSLIDTYIRDPKREFLFNAIETMPYVKKKADWALRWIADRK 180
QY 181 STFGERVVAFPAAVEGVFFSGSFAAIFWLKRGKGLMPGLTFSNELISRDEGLHCDPACLMFQ 240
DB 181 STFGERVVAFPAAVEGVFFSGSFAAIFWLKRGKGLMPGLTFSNELISRDEGLHCDPACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLVLELGS 300
DB 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLVLELGS 300
QY 301 KVFOAENPFDPMENISLEGKTNFFKRVSEYQRFVAVMAETTDNVFTLDADF 351
DB 301 KVFOAENPFDPMENISLEGKTNFFKRVSEYQRFVAVMAETTDNVFTLDADF 351

RESULT 14
US-11-088-686-25
; Sequence 25, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
```

;
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-088-686-25

Query Match
Best Local Similarity 99.4%; Score 1812; DB 7; Length 351;
Matches 349; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDQDERSSDTNESIKNEEPLAKSSRRFVIFPIQYPDIMWKYKQAO 60
Db 1 MGDPERPEAAGLDQDERSSDTNESIKNEEPLAKSSRRFVIFPIQYPDIMWKYKQAO 60

QY 61 ASFWTAAEVDLSKOLPHWNKLKADEKYFISHILAFPAASDGI VNNLVERFSOEVOVPEA 120
Db 61 ASFWTAAEVDLSKOLPHWNKLKADEKYFISHILAFPAASDGI VNNLVERFSOEVOVPEA 120

QY 121 RCFYGFQILLENVHSEMYSLIDITYIRDPKREFLNAIETMPYVKKADWALRWIADRK 180
Db 121 RCFYGFQILLENVHSEMYSLIDITYIRDPKREFLNAIETMPYVKKADWALRWIADRK 180

QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGMLPGLTFSNELISRDEGLHCDPACLMFQ 240
Db 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGMLPGLTFSNELISRDEGLHCDPACLMFQ 240

QY 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLVLVELGFS 300
Db 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLVLVELGFS 300

QY 301 KVFOAENPPDFMENISLEGKTNPFKRVSEYQRFVAVMAETTDNVFTLDADF 351
Db 301 KVFOAENPPDFMENISLEGKTNPFKRVSEYQRFVAVMAETTDNVFTLDADF 351

RESULT 15
PCT-US05-09639-27
; Sequence 27, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-09639-27

Query Match
Best Local Similarity 99.4%; Score 1811; DB 1; Length 351;
Matches 349; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPERPEAAGLDQDERSSDTNESIKNEEPLAKSSRRFVIFPIQYPDIMWKYKQAO 60
Db 1 MGDPERPEAAGLDQDERSSDTNESIKNEEPLAKSSRRFVIFPIQYPDIMWKYKQAO 60

QY 61 ASFWTAAEVDLSKOLPHWNKLKADEKYFISHILAFPAASDGI VNNLVERFSOEVOVPEA 120
Db 61 ASFWTAAEVDLSKOLPHWNKLKADEKYFISHILAFPAASDGI VNNLVERFSOEVOVPEA 120

QY 121 RCFYGFQILLENVHSEMYSLIDITYIRDPKREFLNAIETMPYVKKADWALRWIADRK 180
Db 121 RCFYGFQILLENVHSEMYSLIDITYIRDPKREFLNAIETMPYVKKADWALRWIADRK 180

Db 121 RCFWGFQILLENVHSEMYSLIDITYIRDPKREFLNAIETMPYVKKADWALRWIADRK 180
QY 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGMLPGLTFSNELISRDEGLHCDPACLMFQ 240
Db 181 STFGERVVAFAAVEGVFFSGSFAAIFWLKRGMLPGLTFSNELISRDEGLHCDPACLMFQ 240
QY 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLVLVELGFS 300
Db 241 YLVNKPSEERVREIIVDAVKIEQEFLEALPVGLIGMNCILMKQYIEFVADRLVLVELGFS 300
QY 301 KVFOAENPPDFMENISLEGKTNPFKRVSEYQRFVAVMAETTDNVFTLDADF 351
Db 301 KVFOAENPPDFMENISLEGKTNPFKRVSEYQRFVAVMAETTDNVFTLDADF 351

Search completed: October 30, 2005, 07:02:43
Job time : 45 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 30, 2005, 06:24:21 ; Search time 173.25 Seconds
(without alignments)
14198.102 Million cell updates/sec

Title: US-10-698-228-2

Perfect score: 1846

Sequence: 1 atggcgaccgcaaggcc.....tcaccttgatgcagatttt 1053

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 695266 seqs, 1168006243 residues

Total number of hits satisfying chosen parameters: 13918532

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+_n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10698228/runat_26102005_100609_4045/app.query.fasta_1.2446
-DB=Pending Patents AA Main -QFMT=faetan -SUFFIX=n2p.rapm -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS-human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORW=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USBR=US10698228 @cgn 1 1 523 @runat 26102005 100609_4045
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FCGAPOP=6 -FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: Pending Patents AA Main.*
2: /cgn2_6/ptodata/1/paa/PTUS COMB.pcp.*
3: /cgn2_6/ptodata/1/paa/US06 COMB.pcp.*
4: /cgn2_6/ptodata/1/paa/US07 COMB.pcp.*
5: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
6: /cgn2_6/ptodata/1/paa/US082 COMB.pcp.*
7: /cgn2_6/ptodata/1/paa/US083 COMB.pcp.*
8: /cgn2_6/ptodata/1/paa/US084 COMB.pcp.*
9: /cgn2_6/ptodata/1/paa/US085 COMB.pcp.*
10: /cgn2_6/ptodata/1/paa/US086 COMB.pcp.*
11: /cgn2_6/ptodata/1/paa/US087 COMB.pcp.*
12: /cgn2_6/ptodata/1/paa/US088 COMB.pcp.*
13: /cgn2_6/ptodata/1/paa/US089 COMB.pcp.*
14: /cgn2_6/ptodata/1/paa/US090 COMB.pcp.*
15: /cgn2_6/ptodata/1/paa/US091 COMB.pcp.*
16: /cgn2_6/ptodata/1/paa/US092 COMB.pcp.*
17: /cgn2_6/ptodata/1/paa/US093 COMB.pcp.*
18: /cgn2_6/ptodata/1/paa/US094 COMB.pcp.*
19: /cgn2_6/ptodata/1/paa/US095 COMB.pcp.*
20: /cgn2_6/ptodata/1/paa/US096 COMB.pcp.*
21: /cgn2_6/ptodata/1/paa/US097A COMB.pcp.*
22: /cgn2_6/ptodata/1/paa/US097B COMB.pcp.*
23: /cgn2_6/ptodata/1/paa/US098 COMB.pcp.*
24: /cgn2_6/ptodata/1/paa/US099A COMB.pcp.*
25: /cgn2_6/ptodata/1/paa/US099B COMB.pcp.*
26: /cgn2_6/ptodata/1/paa/US100 COMB.pcp.*
27: /cgn2_6/ptodata/1/paa/US101 COMB.pcp.*
28: /cgn2_6/ptodata/1/paa/US102 COMB.pcp.*

29: /cgn2_6/ptodata/1/paa/US103 COMB.pcp.*
30: /cgn2_6/ptodata/1/paa/US104 COMB.pcp.*
31: /cgn2_6/ptodata/1/paa/US105 COMB.pcp.*
32: /cgn2_6/ptodata/1/paa/US106 COMB.pcp.*
33: /cgn2_6/ptodata/1/paa/US107 COMB.pcp.*
34: /cgn2_6/ptodata/1/paa/US108 COMB.pcp.*
35: /cgn2_6/ptodata/1/paa/US109 COMB.pcp.*
36: /cgn2_6/ptodata/1/paa/US110 COMB.pcp.*
37: /cgn2_6/ptodata/1/paa/US60 COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1821	98.6	351	22	US-09-787-491B-11
2	1821	98.6	351	22	US-09-791-537-77474
3	1821	98.6	351	27	US-10-170-205B-19045
4	1821	98.6	351	32	US-10-698-228-1
5	1821	98.6	351	35	US-10-990-328-7547
6	1821	98.6	351	37	US-60-128-660-7
7	1821	98.6	351	37	US-60-449-629-94
8	1821	98.6	351	37	US-60-505-218-322
9	1821	98.6	366	22	US-09-791-537-27671
10	1817	98.4	351	20	US-09-629-469A-12811
11	1817	98.4	351	22	US-09-791-537-112185
12	1817	98.4	351	35	US-10-917-503-12811
13	1454	78.8	389	1	PCT-US02-10824-143
14	1454	78.8	389	1	PCT-US03-17409-171
15	1454	78.8	389	10	US-08-663-617-33
16	1454	78.8	389	12	US-08-837-599-33
17	1454	78.8	389	27	US-09-791-537-88179
18	1454	78.8	389	27	US-10-170-205B-8018
19	1454	78.8	389	30	US-10-408-765-559
20	1454	78.8	389	30	US-10-408-765A-559
21	1454	78.8	389	32	US-10-698-228-5
22	1454	78.8	389	33	US-10-733-878-457
23	1454	78.8	389	33	US-10-756-149-4899
24	1454	78.8	389	35	US-10-940-774-6016
25	1454	78.8	389	37	US-60-389-987-559
26	1454	78.8	389	37	US-60-412-418-559
27	1454	78.8	389	37	US-60-452-680-20616
28	1454	78.8	389	37	US-60-568-073-984
29	1454	78.8	413	1	PCT-US00-05882-1347
30	1454	78.8	413	24	US-09-925-301-1347
31	1454	78.8	449	35	US-10-990-328-12349
32	1454	78.8	449	35	US-10-990-328-12350
33	1454	78.8	453	1	PCT-US01-04926A-686
34	1454	78.8	453	1	PCT-US01-08631-45646
35	1454	78.8	453	28	US-10-220-335-686
36	1454	78.8	453	28	US-10-290-752-686
37	1454	78.8	453	35	US-10-940-774-7896
38	1443	78.2	390	10	US-08-663-617-35
39	1443	78.2	390	12	US-08-837-599-35
40	1443	78.2	390	22	US-09-791-537-117257
41	1425.5	77.2	386	22	US-09-791-537-114288
42	1425.5	77.2	386	30	US-10-403-571-76
43	1425.5	77.2	386	30	US-10-405-887-76
44	1421	77.0	386	22	US-09-791-537-36076
45	1349.5	73.1	399	22	US-09-791-537-1129

ALIGNMENTS

RESULT 1

US-09-787-491B-11
; Sequence 11, Application US/09787491B
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; TANG, Y. Tom;

APPLICANT: CORLEY, Neil C.; GUEGLER, Karl J.;
APPLICANT: GORONE, Gina A.; PATTERSON, Chandra;
APPLICANT: HILLMAN, Jennifer L.; BAUGHN, Mariah R.;
APPLICANT: LAL, Preeti; AZIMZAI, Yalda;
APPLICANT: YUE, Henry; YANG, Junming
TITLE OF INVENTION: RNA-ASSOCIATED PROTEINS
FILE REFERENCE: PF-0600 USN
CURRENT APPLICATION NUMBER: US/09/787,491B
CURRENT FILING DATE: 2002-11-04
PRIOR APPLICATION NUMBER: PCT/US99/21688
PRIOR FILING DATE: 1999-09-17
PRIOR APPLICATION NUMBER: US 60/128,660
PRIOR FILING DATE: 1999-04-08
PRIOR APPLICATION NUMBER: US 60/069,391
PRIOR FILING DATE: 1998-11-04
PRIOR APPLICATION NUMBER: US 60/183,025
PRIOR FILING DATE: 1998-09-22
PRIOR APPLICATION NUMBER: US 60/155,246
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: US 09/158,720
PRIOR FILING DATE: 1998-09-17
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PERL Program
SEQ ID NO 11
LENGTH: 351
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: 2604449CD1
US-09-787-491B-11

Alignment Scores:

Pred. No.:	1.89e-186	Length:	351
Score:	1821.00	Matches:	351
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	98.65%	Indels:	0
DB:	22	Gaps:	0

US-10-698-228-2 (1-1053) x US-09-787-491B-11 (1-351)

QY	1	ATGGGCGACCGGAAAGCGGAGCGCGGCTGGATCAGGATGAGATCATCTTCA	60
DB	1	MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSer	20
QY	61	GACACCAAGAGTGAATAAGTCAATGAGAGCCACTCTCTAAGAAAGAGTTCTCGC	120
DB	21	AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerArg	40
QY	121	CGTTTGTCTATCTTCCATCCAGTACCCTGATATTGGAAATGTATAACAGGCACAG	180
DB	41	ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln	60
QY	181	GCTTCTCTCTGACACAGAGAGGTCGACTTATCAAGAGTCTCCTCACTGGAACAAG	240
DB	61	AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys	80
QY	241	CTTAAAGCAGATCAGAAGTACTTCTCATCTCTCACATCTTAGCCCTTTTTCGACGCAGTGT	300
DB	81	LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp	100
QY	301	GGAAATGTAAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGTTCCAGAGGCT	360
DB	101	GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnProGluAla	120
QY	361	CGCTGTTTCTATGGCTTTCAATTTCTCATCGAAGTCTTCACTCAGAGATGTACATTTG	420
DB	121	ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu	140
QY	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTTGAA	480
DB	141	LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu	160

QY	481	ACCATGCCCTATGTTAAGAAAAAGACAGATTGGGCCCTTTCGATGATAGCAGATAGAAAA	540
DB	161	ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys	180
QY	541	TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTCTGTAGAAAGAGTTTCTTCTCAGA	600
DB	181	SerThrPheGlyGluArgValAlaPheAlaValGluGlyValPhePheSerGly	200
QY	601	TCCTTTGCTCTATATTCTCGCTAAAGAGAGGTCTTATGCCAGGACTCCTTTTTC	660
DB	201	SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer	220
QY	661	AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGCTGCTTGTGCTGCTGCTGCT	720
DB	221	AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln	240
QY	721	TACTTAGTAATAAGCTTTCAGAAAGAGGTTCAGGAGATCATTTGTGATGCTCTCAAA	780
DB	241	TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys	260
QY	781	ATTGAGCAGCAGCTTTTAAACAGAGCCTTCCAGTTCGCTCATTTGGAATGAATTCATT	840
DB	261	IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle	280
QY	841	TTGATGAACAGTACATTGAGTTTGTAGTCACAGATTACTTGTGGAACCTTGGATTCTCA	900
DB	281	LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer	300
QY	901	AAGGTTTTTCAGGCAGAAAAATCCTTTTGTATTTATGAAAAACATTTCTTTAGAGGAAA	960
DB	301	LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys	320
QY	961	ACAAATTTCTTTCAGAAACAGATTTCAGATATCAGCGTTTTCAGTTTATGCGACAAACC	1020
DB	321	ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr	340
QY	1021	ACAGATAACGCTCTTCACTTGGATGACAGATTTT 1053	
DB	341	ThrAspAsnValPheThrLeuAspAlaAspPhe 351	

RESULT 2

US-09-791-537-77474

; Sequence 77474, Application US/09791537

; GENERAL INFORMATION:

; APPLICANT: Bionomix, Inc.

; APPLICANT: Debe, Derek

; APPLICANT: Danzer, Joseph

; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE

; FILE OF INVENTION: METHODS OF USE THEREOF

; FILE REFERENCE: 261/210

; CURRENT APPLICATION NUMBER: US/09/791,537

; CURRENT FILING DATE: 2001-02-22

; NUMBER OF SEQ ID NOS: 153055

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 77474

; LENGTH: 351

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-791-537-77474

Alignment Scores:

Pred. No.:	1.89e-186	Length:	351
Score:	1821.00	Matches:	351
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	98.65%	Indels:	0
DB:	22	Gaps:	0

US-10-698-228-2 (1-1053) x US-09-791-537-77474 (1-351)

QY	1	ATGGGCGACCGGAAAGCGGAGCGCGGCTGGATCAGGATGAGATCATCTTCA	60
DB			

1	MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer	20
61	GACACCAACGAAGTGAATAAAGTCAAATGAAGAGCCACTCTCTAAGAAGAGAGTGTCTCGC	120
21	AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg	40
121	CGGTTTGTGCATCTTTCCAAATCCAGTACCCTGATATTTCGAAAAATGTATAACAGGCACAG	180
41	ArgPheValIlePheProIleGlnTyrProAspIleTrrPlysMetTyrLysGlnAlaGln	60
181	GCTTCCTTCTCGACAGCAGAGAAGAGTCCACATTATCAAAAGGATCTCCCTCACTTGGAAACAAG	240
61	AlaSerPheTrrPThrAlaGluGluValAspLeuSerLysAspLeuProHisTrrPasnLys	80
241	CTTAAAGCAGATGAGAAGTACTTCATCTCTCACATCTTAGCCCTTTTTCGACCCAGTCAT	300
81	LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp	100
301	GGAAATGTAATGAATAATTTGGTCGAGCGCTTGTAGTCAGGAGGTGCAGGTTCACAGAGCT	360
101	GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla	120
361	CGCTGTTTCTATGGCTTTCAAATCTCATCGAGAATGTTCACTCAGAGATGTCACAGTTTG	420
121	ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu	140
421	CTGATAGACACTTACATCAGAGATCCCMAGAAAAAGGGAATTTTATTTAATGCAATTGAA	480
141	LeuIleAspThrTrrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu	160
481	ACCATGCCCTATGTTAAGAAAAAGCAGATTGGCGCTTGGCATCGATAGCAGATAGAGAAA	540
161	ThrMetProTyrValLysLysLysAlaAspTrrPalaLeuArgTrrPileAlaAspArgLys	180
541	TCTACTTTTGGGAAAGAGTGGTGGCGCTTCTGCTGTAGAAGAGATTTTCTTCTCAGGA	600
181	SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly	200
601	TCTTTTCTGCTATATTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACATTTTCC	660
201	SerPheAlaAlaIlePheTrrPLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer	220
661	AATGAACCTCATCAGCAGAGATGAAGGACTTCACCTGTGACTTGTGCTGCTGATGTTCCAA	720
221	AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln	240
721	TACTTAGTAATAAGCCTTCAGAAAGAAGGGTCAAGGAGATCATTTGTGTGCTGTCAAA	780
241	TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys	260
781	ATTGAGCAGAGATTTTAAACAGAAAGCCTTGCAGTGCCTCATTTGGAATGAATTCGATT	840
261	IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle	280
841	TTGATGAACAACGATCATTTGATTTGTAGCTGACAGATTTACTTGTGGAACTTGGATTTCTCA	900
281	LeuMetLysGlnTrrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer	300
901	AAGGTTTTTCAGGCAGAAAAATCCTTTTGAATTTATGAAAAACATTTCTTTAGAAGGAAAA	960
301	LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys	320
961	ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTCAGCGTTTTCAGTTATGGCAAAACC	1020
321	ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr	340
1021	ACAGATAACGCTTTCACCTTGGATGCAGATTTT	1053
341	ThrAspAsnValPheThrLeuAspAlaAspPhe	351

RESULT 3

US-10-170-205E-19045

; Sequence 19045, Application US/10170205E

; GENERAL INFORMATION:

APPLICANT: ADAMS, Mark

	CLASS OF INVENTION	
; TITLE OF INVENTION:	DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN	
; TITLE OF INVENTION:	CAPTURE AGENTS, AND USES THEREOF	

; TITLE OF INVENTION: CAPTURE AGENTS, AND USES THEREOF
; FILE REFERENCE: CL001381

```

; FILE REFERENCE: CL001361
; CURRENT APPLICATION NUMB

```

; CURRENT AFFILIATION NUMBER: 05/10/110,203
 : CURRENT FILING DATE: 2002-06-13

NUMBER OF SEO ID NOS: 40312

; NUMBER OF SEQ ID NOS: 40312
: SOFTWARE: PatentIn version 1.0.6

```

; SOFTWARE: PALENCIN VERSION 3.2
: SEO ID NO 19045

```

```

; SEQ ID NO 19045
:
: LENGTH: 351

```

```
; LENGTH: 351
; TYPE: pdf
```

TYPE: PRINT
ORGANISM:

ORGANISM: Homo sapiens

US-10-170-205E-19045

Alignment Scores:

Alignment scores:
Pred. No.: 1.89e-186

Field: no.:	1:03e-100	Length:	331
Score:	1821.00	Matches:	351

Score: 1821.00
Percent Similarity: 100.00%
Matches: 100.00%
Conservation: 100.00%

Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Mismatches: 0

Best Local Similarity:	100.00%	Mismatch
Query Match:	98.65%	Indels

Query match:	98.63%	Index:	0
DB:	27	Class:	0

US-10-698-228-2 (1-1053) x US-10-170-205E-19045 (1-351)

1	ATGGCGCAGCCCGGAAAGCCCGGAGCCGCGCTGGATCAGATGAGAGATCATCTTCA	60
Qy		
1	MetGlyAspProGluAtrProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer	20
Db		
61	GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTCTAAGAAGAGTTCCTCG	120
Qy		
21	AspThrAsnGluSerGluLeuIleuLysSerAsnGluGluProLeuLeuArgLysSerArg	40
Db		
121	CGGTTTGTCTATCTTTCCAATCCAGTACCCTCATATTTGGAAAATGTATATAACAGGCACAG	180
Qy		
41	ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln	60
Db		
181	GCTTCTCTTGGACAGCAGAAAGAGTCGACTTATCAAAAGGATCTCCCTCACTGGAAACAAG	240
Qy		
61	AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys	80
Db		
241	CTTAAACGACAGATGAGAAGTACTTCTCATCTCATCTTTAGCCCTTTTTCAGCCAGTGAT	300
Qy		
81	LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp	100
Db		
301	GGAAATCTTAATGAAATTTTCGTGGAGCGCTTTAGTCAGGAGGTCCAGGTTCCAGAGGCT	360
Qy		
101	GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla	120
Db		
361	CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG	420
Qy		
121	ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu	140
Db		
421	CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTGAA	480
Qy		
141	LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu	160
Db		
481	ACCATGCCCTATGTTTAGAAAAAGCAGATTGGCCCTTCCGATGTAGTACAGATAGAAAA	540
Qy		
161	ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys	180
Db		
541	TCTACTTTTGGGAAAGAGTCGTGGCCCTTCTGCTGTAGAAGAGGAGTTTCTTCTCAGGA	600
Qy		
181	SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly	200
Db		
601	TCTTTTGTCTATATTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTTC	660
Qy		
201	SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer	220
Db		
661	AATGAACCTCATCAGCAGATGAAGGACTTCACGTGTGACTTTGCTGCTGATGTTCGAA	720
Qy		
221	AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln	240
Db		

QY 721 TACTTAGTAATAAGCTTTCAGAGAAAGGGTCAGGAGATCATTTGTCATGCTGTCAAA 780
Db 241 TyrLeuValAsnLysProSerGluArgValArgGluIleValAspAlaValLys 260
QY 781 ATTGACGAGGAGTTTAAACAGAGCTTGCAGTTGGCTCATTTGGAATGAATTCAT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAACAGCTACATTCAGTTTGTAGCTGACAGATTACTTGTGAACTTGGATTCTCA 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTCAGCAGAAATCCTTTGATTTATGGAACAATTTCTTTAGAGGAAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluLys 320
QY 961 ACAATTTCTTTCAGAGAAAGCTTTCAGAGTATCAGCGTTTTCAGATTATGCGAAGAAC 1020
Db 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAAGCTTTCACCTTGGATGCAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 4

US-10-698-228-1
; Sequence 1, Application US/10698228
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; PRIOR FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 1
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-1

Alignment Scores:
Pred. No.: 1.89e-186 Length: 351
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.65% Indels: 0
DB: 32 Gaps: 0

US-10-698-228-2 (1-1053) x US-10-698-228-1 (1-351)

QY 1 ATGGGCGCCCGAAGGCGGAGCGCGGCTCGATCAGGATCAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAAGAAAGTAAAGTCAAAATGAAGAGCCACTCTCAAGAAAGAGTTCTCGC 120
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
QY 121 CGGTTTGTTCATTTCCATCCAGTACCTGTATTTTGAAGAAATGATTAACAGGACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCCTTTCGACAGACAGAGGTCGACTATCAAGAGTCTCCCTCACTGGAACAG 240

Db 61 AlaSerPheTrpThrAlaGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAAGCAGATCAGAGTACTTCTCTCATCTCAGATCTTAGCCTTTTTCAGCAGCAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp 100
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTCGAGGTTCCAGAGGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATGCTTTCAAAATTCATCATCGAGAAATGTTCTCATCAGAGATGATACAGTTG 420
Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTCATAGACACTTACATCAGATCCCAAGAAAGGCAATTTTATTATTAATGCAATTCGA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTGGCATGGATAGCAGATAGAAAA 540
Db 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTGTCTGCTGTAGAGGAGTTTCTTCTCAGA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaValGluGluValPhePheSerGly 200
QY 601 TCTTTTGTCTATATCTGCTTAAAGAACAGAGGCTTATGCCAGGACTCACATTTTTC 660
Db 201 SerPheAlaAlaIlePheTrpLeuLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACCTCATCAGCAGATGAAGGACTTCACGTGCTGCTTGTCTGCTGCTGCTGCTCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAATAAGCTTTCAGAGAAAGGGTCAGGAGATCATTTGTCATGCTGTCAAA 780
Db 241 TyrLeuValAsnLysProSerGluArgValArgGluIleIleValAspAlaValLys 260
QY 781 ATTGACGAGGAGTTTAAACAGAGCTTGCAGTTGGCTCATTTGGAATGAATTCAT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAACAGCTACATTCAGTTTGTAGCTGACAGATTACTTGTGAACTTGGATTCTCA 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTCAGCAGAAATCCTTTGATTTATGGAACAATTTCTTTAGAGGAAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluLys 320
QY 961 ACAATTTCTTTCAGAGAAAGCTTTCAGAGTATCAGCGTTTTCAGATTATGCGAAGAAC 1020
Db 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAAGCTTTCACCTTGGATGCAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 5

US-10-990-328-7547
; Sequence 7547, Application US/10990328
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; TITLE OF INVENTION: USSES THEREOF
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 7547

```
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-990-328-7547

Alignment Scores:
Pred. No.: 1.89e-186 Length: 351
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.65% Indels: 0
DB: 35 Gaps: 0

US-10-698-228-2 (1-1053) x US-10-990-328-7547 (1-351)
QY 1 ATGGCGCGCCGGAAGCGCGGCGCGGCTGGATCAGGATGAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAACGAAAGTAAATAAGTCAAAATGAAGAGCACTCTAAGAAAGAGTCTCGC 120
Db 21 AspThrAsnGluSerGluLeuLeuSerAsnGluGluProLeuLeuArgLysSerArg 40
QY 121 CGTTTGTTCATCTTCCAAATCCAGTACCTCGATATTTGGAATGTATAACAGGACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCTCTTGCACAGCAGAGAGGTCGACTTATCAAGAGTCTCCTCAGTGGAAACAG 240
Db 61 AlaSerPheTrpThrAlaGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAAGCAGATGAGAAGTCAATCTCATCTCTACATCTTAGCCTTTTTCAGCAGTGAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp 100
QY 301 GGAATGTAAATGAAAATTTGGTGAGCGCTTTAGTCAGGAGTGCGAGGTTCCAGAGGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTCTATGCGCTTCAAAATCTCATCGAGAAATGTTCTACATGAGATGATGATGAT 420
Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATCAATTGAA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCGCTATGTTAAAGAAAAAGCAGATTCGGGCTTGGGATGGATGAGATGAGATAA 540
Db 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTGGGAAAGAGTGTGGCTTTGCTGTAGAGAGGATTTTCTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTCTATATTCTGGCTAAGAGAGAGCTCTTATGCCAGGACTCACTTTTCC 660
Db 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTGTGATGTTGCCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAAATAGCCCTTCAAGAAAGGGTCAGGAGATCATTTGTTGATCTCTCAAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
QY 781 ATTGAGCAGGAGTTTAAACAGAACCTTGCAGTTGGCTCATTTGGAATGATTCATT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAACAGTACATTGAGTTTCTAGCTGACAGATTACTTTGTGAACTTGATTTCTCA 900
Db 841 TTTTCTGTTCTTCTGCTTTTCAATTCATCTCAGAAATGTTCTCACTCAGAGATGTACAGTTG 920

RESULT 6
US-60-128-660-7
; Sequence 7, Application US/60128660
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Tang, Y. Tom
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Yang, Junming
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: RNA-ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0682 P
; CURRENT APPLICATION NUMBER: US/60/128,660
; CURRENT FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PERL Program
; SEQ ID NO 7
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 2604449
US-60-128-660-7

Alignment Scores:
Pred. No.: 1.89e-186 Length: 351
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.65% Indels: 0
DB: 37 Gaps: 0

US-10-698-228-2 (1-1053) x US-60-128-660-7 (1-351)
QY 1 ATGGCGCGCCGGAAGCGCGGCGGCTGGATCAGGATGAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAACGAAAGTAAATAAGTCAAAATGAAGAGCACTCTAAGAAAGAGTCTCGC 120
Db 21 AspThrAsnGluSerGluLeuLeuSerAsnGluGluProLeuLeuArgLysSerArg 40
QY 121 CGTTTGTTCATCTTCCAAATCCAGTACCTCGATATTTGGAATGTATAACAGGACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCTCTTGCACAGCAGAGAGGTCGACTTATCAAGAGTCTCCTCAGTGGAAACAG 240
Db 61 AlaSerPheTrpThrAlaGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAAGCAGATGAGAAGTCAATCTCATCTCTACATCTTAGCCTTTTTCAGCAGTGAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp 100
QY 301 GGAATGTAAATGAAAATTTGGTGAGCGCTTTAGTCAGGAGTGCGAGGTTCCAGAGGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTCTATGCGCTTCAAAATCTCATCGAGAAATGTTCTACATGAGATGATGATGAT 420
Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTAATCAATTGAA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCGCTATGTTAAAGAAAAAGCAGATTCGGGCTTGGGATGGATGAGATGAGATAA 540
Db 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTGGGAAAGAGTGTGGCTTTGCTGTAGAGAGGATTTTCTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTCTATATTCTGGCTAAGAGAGAGCTCTTATGCCAGGACTCACTTTTCC 660
Db 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTGTGATGTTGCCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAAATAGCCCTTCAAGAAAGGGTCAGGAGATCATTTGTTGATCTCTCAAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
QY 781 ATTGAGCAGGAGTTTAAACAGAACCTTGCAGTTGGCTCATTTGGAATGATTCATT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAACAGTACATTGAGTTTCTAGCTGACAGATTACTTTGTGAACTTGATTTCTCA 900
Db 841 TTTTCTGTTCTTCTGCTTTTCAATTCATCTCAGAAATGTTCTCACTCAGAGATGTACAGTTG 920
```

```
Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGATTTTATTTAATCAATTGAA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysIysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAGAAAACAGATTCGGCTTGCAGTGTAGATGATAGCAGATAGAAA 540
Db 161 ThrMetProTyrValIysIysAlaAspThrPheAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGGAAGAGTGTGGCTTTGCTGCTGTAGAGAGTGTCTTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTCTATATTCTGGCTTAAGAGAGAGTCTTATGCTCAGGACTCACTTTTCC 660
Db 201 SerPheAlaAlaIlePheTrpLeuLysIysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGGCTGATGTTCCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAATAAGCCTTTCAGAGAAGGGTCAGGAGATCATTTGTTGATGCTGTCAAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleValAspAlaValLys 260
QY 781 ATTGAGCAGGAGTATTAACAGAACGCTTGCAGATATCAGCGTTTGCAGTATGCGAATTCAT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAACACTACATTCAGTTTCTAGCTGACAGATTAATGCGAATTCGATTCTCA 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTTCAGCAGAAAATCCTTTTGATTTTATGGAACAATTTCTTTAGAAGGAAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluLys 320
QY 961 ACAAAATTTCTTTGAGAACAGTTTCAGAGTATCAGCGTTTGCAGTATGCGAAGAAC 1020
Db 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAAGCTTTCACCTTGGATCAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
```

RESULT 7

```
; Sequence 94, Application US/60449629
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CANCER, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001449
; CURRENT APPLICATION NUMBER: US/60/449,629
; CURRENT FILING DATE: 2003-02-04
; NUMBER OF SEQ ID NOS: 2712
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 94
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-449-629-94
```

Alignment Scores:

Pred. No.:	1,89e-186	Length:	351
Score:	1821.00	Matches:	351
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	98.65%	Indels:	0
DB:	37	Gaps:	0

```
US-10-698-228-2 (1-1053) x US-60-449-629-94 (1-351)
QY 1 ATGGGCGACCCGAAAGCGCGGAGCGGCTCGATCAGATGAGAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAACGAAGTGAAATAAAGTCAAATGAAGAGCCACTCTCTAAGAAAGATTCTCGC 120
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
QY 121 CGTTTGTTCATCTTCCAATCCAGTACCCTGATATTTGGAAATGATATAACAGGCACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCCTCTCTGACACAGAGAGTGCAGTCTTATCAAGAGATCTCCCTCACTGTGAACAAG 240
Db 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCATCATCTTAGCCCTTTTTCAGCCAGTGAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY 301 GGAATTTGTAATAAATAAATTTGTGGAGCGCTTTAGTCAGAGGTGCAAGTTCAGAGGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTCTATGGCTTCAAAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTCATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTCGA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAGAAAACAGAGATTTGGGCTTGCAGTGTAGATGAGATAGAAAA 540
Db 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAGAGTGTGGCTTTCCTGCTGTAAGAGGATTTCTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTCTATATTCTGGCTTAAAGAGAGAGTCTTATGCGAGACTCACATTTTTC 660
Db 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAATAAGCCTTTCAGAGAAGGGTCAGGAGATCATTTGTTGATGCTGTCAAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleValAspAlaValLys 260
QY 781 ATTGAGCAGGAGTATTTAAGAAAGCTTGCAGAGTTCAGGCTCATTTGGAATGAATTCATT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAACACTACATTCAGTTTGTAGCTGACAGATTAATGCGAATTCGATTCTCA 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTTCAGCAGAAAATCCTTTGATTTTATGGAACAATTTCTTTAGAAGGAAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluLys 320
QY 961 ACAAAATTTCTTTGAGAACAGTTTCAGAGTATCAGCGTTTGCAGTATGCGAAGAAC 1020
Db 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAAGCTTTCACCTTGGATCAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
```

RESULT 8

US-60-505-218-322
 ; Sequence 322, Application US/60505218
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; FILE REFERENCE: CL001482
 ; CURRENT APPLICATION NUMBER: US/60/505,218
 ; CURRENT FILING DATE: 2003-09-24
 ; NUMBER OF SEQ ID NOS: 22507
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 322
 ; LENGTH: 351
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-60-505-218-322

Alignment Scores:
 Pred. No.: 1,898-186 Length: 351
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 98.65% Indels: 0
 DB: 37 Gaps: 0

US-10-698-228-2 (1-1053) x US-60-505-218-322 (1-351)

QY 1 ATGGCGAGCCCGAAGAGCGCGCGCTGGATCAGAGATCATCTTCA 60
 Db 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
 QY 61 GACACACGAAAGTGAATAAGTCAATGAAGACCACTCTTAAGAAGAGTTCTCC 120
 Db 21 AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerArg 40
 QY 121 CGTTTGTTCATCTTCCCAATCCAGTACCTGATATTGGAAATGTATAACAGGCACAG 180
 Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
 QY 181 GCTCTCTTCGACAGCAGAGAGGTGCGACTTATCAAGGATCTCCCTCAGTGGACACAG 240
 Db 61 AlaSerPheTrpThrAlaGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
 QY 241 CTTAAGCAGATGAGAGTACTTCACTCTCATCTTACATCTTAGCCCTTTTTCAGCCAGTGTAT 300
 Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
 QY 301 GGAATTGTAATGAAATTTGGTGGAGCGCTTACTCAGGAGTCCAGGTTCCAGGCT 360
 Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 QY 361 CGCTGTTTCTATGGCTTTTCAATCTCATCGAGAATGTTTCACTCAGAGATGACAGTTTG 420
 Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 QY 421 CTGATAGACATTCATCAGAGATCCCAAGAAAGGGAATTTTATTAATCAATTGAA 480
 Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 QY 481 ACCATGCCCTATGTTAAGAAAGACAGATTGGGCTTCCGATGATAGCAGATAGAAA 540
 Db 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
 QY 541 TCTACTTTGGGAAAGAGTGGTGGCTTGTCTGTAGAGAGGATTTTCTTCTCAGGA 600
 Db 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
 QY 601 TCTTTTGTCTATATTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTCC 660
 Db 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220

QY 661 AATCAACTCATCAGCAGAGATGAAGGACTTCACTGTGTGCTTGTGCTGTGCTGCTCA 720
 Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 QY 721 TACTTAGTAATAAGCCTTCAGAAAGAGGTGAGGAGATCATTTGATGCTGCTCA 780
 Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
 QY 781 ATTGAGCAGGAGTATTAACAGAGAGCTTGCCAGTTGGCCCTCATTTGGAATGAATTCATT 840
 Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 QY 841 TTGATCAACACAGTACATTTAGTTTGTAGCTGACAGATTACTTGTGAACTTGGATTCTCA 900
 Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 QY 901 AAGTTTTTCAGGAGAAATCTTTTGTATTTTATGAAACATTTCTTTTGAAGAAAA 960
 Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 QY 961 ACAAAATTTCTTCAGAACAGATTTCAGAGTATCAGCGTTTTCGAGTTTATGCGACANAACC 1020
 Db 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 QY 1021 ACAGATAACGCTTTCACCTTGGATGACAGATTTT 1053
 Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 9

US-09-791-537-27671
 ; Sequence 27671, Application US/09791537
 ; GENERAL INFORMATION:
 ; APPLICANT: Bionomix, Inc.
 ; APPLICANT: Debe, Derek
 ; APPLICANT: Danzer, Joseph
 ; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
 ; FILE REFERENCE: 261/210
 ; CURRENT APPLICATION NUMBER: US/09/791,537
 ; CURRENT FILING DATE: 2001-02-22
 ; NUMBER OF SEQ ID NOS: 153055
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 27671
 ; LENGTH: 366
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-791-537-27671

Alignment Scores:

Pred. No.: 1,928-186 Length: 366
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 98.65% Indels: 0
 DB: 22 Gaps: 0

US-10-698-228-2 (1-1053) x US-09-791-537-27671 (1-366)

QY 1 ATGGCGAGCCCGAAGAGCGCGCGCTGGATCAGAGATCAGATCATCTTCA 60
 Db 16 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 35
 QY 61 GACACCAAGAAAGTGAATAAGTCAATGAAGACCACTCTTAAGAAGAGTTCTCGC 120
 Db 36 AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerSerArg 55
 QY 121 CGTTTGTTCATCTTCCAAATCCAGTACCTGATATTGGAAATGTATAACAGGCACAG 180
 Db 56 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 75
 QY 181 GCTTCTCTTCGACAGCAGAGAGGTGCGACTTATCAAGAGATCTCCCTCAGTGGACACAG 240
 Db 76 AlaSerPheTrpThrAlaGluValAspLeuSerLysAspLeuProHisTrpAsnLys 95

```
QY 241 CTTAAAGCAGATGAGAAGTACTTCTCTCATCTTAGCCCTTTTTCAGCCAGTGAT 300
Db 96 LeuLysAlaAspGluLysTyrPheIleSerHisLeuAlaPhePheAlaSerAsp 115
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTCCAGAGGCT 360
Db 116 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValProGluAla 135
QY 361 CGCTGTTTCTATGGCTTTCAAAATTCATCGAGATGTTCACTCAGAGATGACAGTTTG 420
Db 136 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 155
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTAAATGCAATTTGAA 480
Db 156 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 175
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTCGGGCTTTCGATGGATAGCAGATAGAAA 540
Db 176 ThrMetProTyrValLysLysAlaAspTTPAlaLeuArgTTPileAlaAspArgLys 195
QY 541 TCTACTTTTGGGAAAGAGTGTGGCTTTCCTGCTGTAGAGAGTTCCTTCTCAGGA 600
Db 196 SerThrPheGlyArgValValAlaPheAlaValGluGlyValPhePheSerGly 215
QY 601 TCTTTTGTCTATATTCTGGCTTAAAGAGAGAGGTCTTATCCAGAGCTCACTTTTCC 660
Db 216 SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 235
QY 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 236 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 255
QY 721 TACTTAGTAATAAGCTTCAGAGAAGGGTTCAGGAGATCATTTGTTGATGCTGTCAAA 780
Db 256 TyrLeuValAsnLysProSerGluGluArgValArgGluIleValAlaValLys 275
QY 781 ATTGAGCAGGAGTATTAACAGAAGCTTGCAGTTGGCTTCACTTGGATGATGATTCATT 840
Db 276 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 295
QY 841 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTTACTTGTGGAATTCGATTTCTCA 900
Db 296 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 315
QY 901 AAGTTTTTTCAGCGAGAAATCCTTTTGTATTTATGGAACAACTTCTTTAGAGAGAAA 960
Db 316 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 335
QY 961 ACNAATTTCTTGAGAAAGAGTTTCAGAGTATCAGCGTTTTCAGATTTATGCGAGAAACC 1020
Db 336 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 355
QY 1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053
Db 356 ThrAspAsnValPheThrLeuAspAlaAspPhe 366
```

RESULT 10

US-09-629-469A-12811

; Sequence 12811, Application US/09629469A

; GENERAL INFORMATION:

; APPLICANT: OTA, TOSHIO

; APPLICANT: ISOGAI, TAKAO

; APPLICANT: NISHIKAWA, TETSUO

; APPLICANT: HAYASHI, KOJI

; APPLICANT: SAITO, KAORU

; APPLICANT: YAMAMOTO, JUNICHI

; APPLICANT: ISHII, SHIZUKO

; APPLICANT: SUGIYAMA, TOMOYASU

; APPLICANT: WAKAMATSU, AI

; APPLICANT: NAGAI, KEIICHI

; APPLICANT: OTSUKI, TETSUJI

; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH CDNA AND THEIR USE

```
; FILE REFERENCE: 084335/0123
; CURRENT APPLICATION NUMBER: US/09/629,469A
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: JP 1999-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 1999-300253
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/183,322
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 19025
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12811
; LENGTH: 351
; TYPE: PRI
; ORGANISM: Homo sapiens
; US-09-629-469A-12811
```

Alignment Scores:

Pred. No.:	5,11e-186	Length:	351
Score:	1817.00	Matches:	350
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.72%	Mismatches:	0
Query Match:	98.43%	Indels:	0
DB:	20	Gaps:	0

US-10-698-228-2 (1-1053) x US-09-629-469A-12811 (1-351)

```
QY 1' ATGGGCGACCCCGAAGGCGGAGCGCGGCTGATCAGATGAGAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAAGAAAGTGAATTAAGTCAATTAAGAGAGCCATCTCTAAGAAAGATCTCGC 120
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerSerArg 40
QY 121 CGGTTTTCATCTTCCCAATCCAGTACCTGATATTTGGAAATGTAACAGGCACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCCTTCTGGACAGACAGAGGTCGACTTATCAAGAGATCTCCCTCACTGGAACAAG 240
Db 61 AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
QY 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACAATCTTAGCCTTTTTCAGCCAGTGAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp 100
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGGAGGTCCAGAGGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATGGCTTTCAAAATTCATCAGAGATGTTCTACTCAGAGATGTACAGTTTG 420
Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTAAATGCAATTTGAA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCTTGGATGGATGAGATAGAAAA 540
Db 161 ThrMetProTyrValLysLysLysAlaAspTTPAlaLeuArgTTPileAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAGAGTGTGGCTTTCCTGCTGTAGAGAGTTCCTTCTCAGGA 600
```

Db 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTCTATATTCTGGCTAAAGAGAGAGGTCTTATGCGAGGACTCACTTTTCC 660
Db 201 SerPheAlaAlaIlePheTrpLeuIleYsYsArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACATCATCAGCAGAGATGAAGGACTTCTACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAATAAGCCTTCAGAAAGAGGTTCAGGAGATCATTTGCTGCTCAAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValAlaGluIleValAlaValLys 260
QY 781 ATTGACGAGGAGTCTTAAACAGAGCCTTGCAGTGGCTCATTTGGAATGAATGCATT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyValAsnCysIle 280
QY 841 TTGATGAACACATGATGATTGCTGAGTGTGCTGACAGATTAATTTGGAACCTTGGATTCTCA 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGTTTTTCAGGAGAGAAATCTTTGATTTTATGGAATAATTTCTTTAGAGAGAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAATTTCTTTGAGAACGAGTTTCAGAGTATCAGCGTTTTCAGTGTATGCGAGAAC 1020
Db 321 ThrAsnPheGluGlyArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 11

US-09-791-537-112185
; Sequence 112185, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Blonoxim, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 112185
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-112185

Alignment Scores:
Align. No.: 5,11e-186 Length: 351
Score: 1817.00 Matches: 350
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.72% Mismatches: 0
Query Match: 98.43% Indels: 0
DB: 22 Gaps: 0

US-10-698-228-2 (1-1053) x US-09-791-537-112185 (1-351)

QY 1 ATGGCGCAGCCGGAAGCCGGAAGCGCGCGGCTGGATCAGATGAGAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAACGAAGTGAATAAGTCAATGAAGAGCCATCTCTAAGAAAGAGTTCTCGC 120
Db 21 AspThrAsnGluSerGluIleYsSerAsnGluGluProLeuLeuArgLysSerArg 40
QY 121 CGTTTGTCTCTTCCATTCAGTACCTGATATTTGGAAATGTATTAACAGGCACAG 180

RESULT 12

US-10-917-503-12811
; Sequence 12811, Application US/10917503
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: SAITO, KAORU
; APPLICANT: YAMAMOTO, JUNICHI

Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCTCTTGGACAGCAGAGAGGTGCTTATCAAGGATCTCCCTCACTGGAACAAG 240
Db 61 AlaserPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAAGCAGATCAGAAAGTACTTCTCATCTCTCACATCTTAGCCTTTTTCAGCAGCAGTGTAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAspSerAsp 100
QY 301 GGAATTTGTAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATGCTTCAAAATCTCATCGAGATGTTTCTACTCAGAGATGTACAGTTTG 420
Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTCGA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTTATGTTAAGAAAGAGATTTGGCGCTTGCAGTGCATGATGATAGATAAAA 540
Db 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAGAGGTGGCTTTGCTGCTGTAGAGAGGATTTTCTTCTCAGA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTCTATATTCTGGCTAAAGAGAGAGGTCTTATGCCAGACTCACATTTTCC 660
Db 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACATCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTCTGCTCATGTTCCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAATAAGCCTTCAGAAAGAGGTTCAGGAGATCATTTGATGCTGTCAAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAlaValLys 260
QY 781 ATTGACGAGGATTTTAAACAGAGCCTTCCAGTGGCTTCATTTGGCTCATTTGGAATGCATT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyValAsnCysIle 280
QY 841 TTGATCAAAACAGTACATGAGTTTGTAGCTGACAGATTAATTTGGAACCTTGGATTCTCA 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGTTTTTTCAGGCAGAAATCTTTTGTATTTTATGGAATAATTTCTTTAGAGAGAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAATTTCTTTGAGAACGAGTTTTCAGAGTATCAGCGTTTTCAGTGTATGCGAGAAC 1020
Db 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

APPLICANT: ISHII, SHIZUKO
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: WAKAMATSU, AI
APPLICANT: NAGAI, KEIICHI
APPLICANT: OTSUKI, TETSUJI
TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE
FILE REFERENCE: 084335/0123
CURRENT APPLICATION NUMBER: US/10/917,503
CURRENT FILING DATE: 2004-08-13
PRIOR APPLICATION NUMBER: US/09/629,469
PRIOR FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: JP 1999-248036
PRIOR FILING DATE: 1999-07-29
PRIOR APPLICATION NUMBER: JP 1999-300253
PRIOR FILING DATE: 1999-08-27
PRIOR APPLICATION NUMBER: JP 2000-118776
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: JP 2000-183767
PRIOR FILING DATE: 2000-05-02
PRIOR APPLICATION NUMBER: JP 2000-241899
PRIOR FILING DATE: 2000-06-09
PRIOR APPLICATION NUMBER: 60/159,590
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 60/183,322
PRIOR FILING DATE: 2000-02-17
NUMBER OF SEQ ID NOS: 19025
SOFTWARE: Patent in Ver. 2.1
SEQ ID NO 12811
LENGTH: 351
TYPE: PRT
ORGANISM: Homo sapiens
US-10-917-503-12811

Alignment Scores:
Pred. No.: 5,11e-186 Length: 351
Score: 1817.00 Matches: 350
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.72% Mismatches: 0
Query Match: 98.43% Indels: 0
DB: 35 Gaps: 0

US-10-698-228-2 (1-1053) x US-10-917-503-12811 (1-351)

QY 1 ATGGCGACCCGGAAGCCGGAAGCCGGCTGGATCAGATCAGATCATCTTCA 60
DB 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAACGAAAGTGAATAAGTCAATGAAGACCACTCTTAAGAAAGAGTTCTCGC 120
DB 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
QY 121 CGTTTGTTCATCTTCCATCCAGTACCTCTGATATTGTGAAATGATATAACAGGCACAG 180
DB 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCCTTCTGGACAGCAGACAGGTGCGACTTATCAAGAGTCTCCCTCAGTGGACAG 240
DB 61 AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
QY 241 CTTAAGCAGATGAGAAGTACTTCTCTCATCTTACATCTTAGCCCTTTTTCAGCCAGTGAT 300
DB 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAlaSerAsp 100
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGTTTCCAGAGGCT 360
DB 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTCTATGGCTTTCAAAATTCATCGAGATGTTCTCAGAGATGTACAGTTTG 420
DB 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACACTTACATCAGACATCCCAAGAAAGGAATTTTATTTAATGCAATTGAA 480

DB 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATCTTAAAGAAAAGCAGATGGCCCTTGCATGATGATGATGATGATAA 540
DB 161 ThrMetProTyrValLysLysLysAlaAspTyrPalaLeuArgTyrPalaAspArgLys 180
QY 541 TCTACTTTTGGGAAAAGAGTGGTGGCTTCTGCTGCTAGAGGAGTCTTCTTCTCAGGA 600
DB 181 SerThrPheGlyGluArgValValAlaPheAlaValGluGluValPhePheSerGly 200
QY 601 TCTTTTGTCTATATCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACATTTTTC 660
DB 201 SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACGTGACCTTGTCTGCTGCTGCTGCTCAA 720
DB 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAAATAAGCCTTCAGAAAGAAAGGCTCAGGAGAGATCATTTGTGCTGCTCAA 780
DB 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
QY 781 ATTGACAGAGAGTCTTAAACAGAGCTTCCAGTTCGCCCTCATTTGGCAATGCAATTCAT 840
DB 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyValAsnCysIle 280
QY 841 TTGATCAACACAGTACATTTAGTTTGTAGCTGACAGATTTACTTGTGGAATTTGGATTCTCA 900
DB 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTTCAGGACAGAAATCCTTTTGTATTTTATGAAACATTTCTTTAGAAGGAAAA 960
DB 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAATTTCTTTCAGAAACAGTTCAGAGTATCAGCGTTTTCAGCTTTTTCAGCTTATGCGAGAAACC 1020
DB 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGTCTTCCCTTGGATGATGATTTT 1053
DB 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 13
PCT-US02-10824-143
Sequence 143, Application PC/TUS0210824
GENERAL INFORMATION:
APPLICANT: Origene Technologies
TITLE OF INVENTION: Prostate Cancer Expression Profiles
FILE REFERENCE: 9U 206 PCT
CURRENT APPLICATION NUMBER: PCT/US02/10824
CURRENT FILING DATE: 2002-04-08
PRIOR APPLICATION NUMBER: US 60/281,732
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: US 60/281,731
PRIOR FILING DATE: 2001-04-06
NUMBER OF SEQ ID NOS: 211
SOFTWARE: Patent in version 3.1
SEQ ID NO 143
LENGTH: 389
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US02-10824-143

Alignment Scores:
Pred. No.: 7,67e-147 Length: 389
Score: 1454.00 Matches: 269
Percent Similarity: 94.08% Conservative: 33
Best Local Similarity: 83.80% Mismatches: 19
Query Match: 78.76% Indels: 0
DB: 1 Gaps: 0

US-10-698-228-2 (1-1053) x PCT-US02-10824-143 (1-389)

QY	91	GAAGAGCCACTCTTAAGAAAGAGTTCTCGCGGTTTGTCACTTTCCAAATCCAGTACCCT	150
DB	69	AspGluProLeuLeuArgGluAsnProArgPheValIlePheProIleGluTyrHis	88
QY	151	GATATTTGGAAATGTATAACAGCAGCAGGCTTCCTCTCGACAGAGAGAGGTGCAC	210
DB	89	AspIleTrpGlnMetTyrLysLysAlaGluAsnSerPheTrpThrAlaGluGluValAsp	108
QY	211	TTATCAAGAGTATCCCTCACTGGAAACAAGCTTAAAGCAGATGAGAAGTACTTTCATCTCT	270
DB	109	LeuSerLysAspIleGlnHisTrpGluSerLeuLysProGluGluArgTyrPheIleSer	128
QY	271	CACATCTTAGCCCTTTTTCGAGCCAGTATGTAATGAAATTTGGTGGAGCGC	330
DB	129	HisValLeuAlaPhePheAlaAlaSerAspGlyIleValAsnGluAsnLeuValGluArg	148
QY	331	TTTAGTCAGGAGGTGCAGGTTCCAGAGGCTCGCTGTTTCTATGCGCTTTCAAATTCCTCATC	390
DB	149	PheSerGlnGluValGlnIleThrGluAlaArgCysPheTyrGlyPheGlnIleAlaMet	168
QY	391	GAGAAATGTTCACTCAGAGATGTACAGTTTGCTGTATAGACACTTACATCAGAGATCCCAAG	450
DB	169	GluAsnIleHisSerGluMetTyrSerLeuLeuIleAspThrTyrIleLysAspProLys	188
QY	451	AAAGGCAATTTTATTAAATGCAATTAACCAATGCCCTATGTTAAGAAAAAGCAGAT	510
DB	189	GluArgGluPheLeuPheAsnAlaIleGluThrMetProCysValLysLysLysAlaAsp	208
QY	511	TGGCGCTTGCGATGATAGCAGATAGAAAATCTACTTTGGGAAACAGGTGGCGCTTT	570
DB	209	TrpAlaLeuArgTrpIleGlyAspLysGluAlaThrTyrGlyGluArgValAlaPhe	228
QY	571	GCTGCTGTAGAAGAGTTTCTTCTCAGGATCTTTTCTGCTGCTATATTTCTGGCTAAAGAG	630
DB	229	AlaAlaValGluGlyIlePhePheSerGlySerPheAlaSerIlePheTrpLeuLysLys	248
QY	631	AGAGGCTTATGCCAGGACTCACTTTTCCAAATCAATCATCAGCAGAGATGAAGACTT	690
DB	249	ArgGlyLeuMetProGlyLeuThrPheSerAsnGluLeuIleSerArgAspGluGlyLeu	268
QY	691	CACGTGACATTTGCTGCTGATGTTCCAAATCTAGTAATTAAGCCTTCAGAGAAAGG	750
DB	269	HisCysAspPheAlaCysLeuMetPheLysHisLeuValHisLysProSerGluGluArg	288
QY	751	GTCAAGGAGATCATTTGTGTATGCTGTCAAAATTTAGCAGCAGGTTTTTAAACAGAGCCTTG	810
DB	289	ValArgGluIleIleIleAsnAlaValArgIleGluGlnGluPheLeuThrGluAlaLeu	308
QY	811	CCAGTTGGCGCTCATTTGGAATGAATTCATTTTGATGAAACAGTACATTTGAGTTGTAGCT	870
DB	309	ProValLysLeuIleGlyMetAsnCysThrLeuMetLysGlnTyrIleGluPheValAla	328
QY	871	GACAGATTACTTTGGAACCTTGGATTCCTCAAGGTTTTTTCAGGCAGAAAAATCCCTTTTCAT	930
DB	329	AspArgLeuMetLeuGluLeuGlyPheSerLysValPheArgValGluAsnProPheAsp	348
QY	931	TTTTATGGAACAATTTCTTTAGAGGAAAAACAAATTTCTTTTGAGAAACAGAGTTTCAGAG	990
DB	349	PheMetGluAsnIleSerLeuGluGlyLysThrAsnPhePheGluLysArgValGlyGlu	368
QY	991	TATCAGCGTTTTGCAGTTATGGCAGAAACACACAGATAACCGTCTTTCACCTTTGGATGTCAGAT	1050
DB	369	TyrGlnArgMetGlyValMetSerSerProThrGluAsnSerPheThrLeuAspAlaAsp	388
QY	1051	TTT 1053	
DB	389	Phe 389	

RESULT 14

RESOUR 14
PCT-US03-17409-171

ECI-0303-17409-171
; Sequence 171, Application PC/TUS0317409

: GENERAL INFORMATION:

; APPLICANT: INCYTE CORPORATION.

Qy 1051 TTT 1053
Db |||
389 Phe 389

Search completed: October 30, 2005, 06:47:55
Job time : 182.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 30, 2005, 06:28:55 ; Search time 36 Seconds
(without alignments)

13219.420 Million cell updates/sec

Title: US-10-698-228-2

Perfect score: 1846

Sequence: 1 atggcgccggaaggcc.....tcaccttgatcgagatttt 1053

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 972358 seqs, 225972999 residues

Total number of hits satisfying chosen parameters: 1944716

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+_n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10698228/runat_26102005_100610_4069/app_query.fasta_1.2446
-DB=Pending Patents AA New -SUFFIX=n2p.rapn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40 cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10698228 @cgn 1 1 22 @runat_26102005_100610_4069
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending Patents AA New:

1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US11_NEW_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1821	98.6	351	1	PCT-US05-09639-15
2	1821	98.6	351	1	PCT-US05-21650-34
3	1821	98.6	351	6	US-10-990-328A-7547
4	1821	98.6	351	7	US-11-088-686-15
5	1817	98.4	351	1	PCT-US05-09639-17
6	1817	98.4	351	7	US-11-088-686-17
7	1816	98.4	351	1	PCT-US05-09639-19
8	1816	98.4	351	7	US-11-088-686-19
9	1813	98.2	351	1	PCT-US05-09639-21
10	1813	98.2	351	7	US-11-088-686-21
					Sequence 15, Appl
					Sequence 34, Appl
					Sequence 7547, Ap
					Sequence 15, Appl
					Sequence 17, Appl
					Sequence 19, Appl
					Sequence 19, Appl
					Sequence 21, Appl
					Sequence 21, Appl

11	1812	98.2	351	1	PCT-US05-09639-23	Sequence 23, Appl
12	1812	98.2	351	1	PCT-US05-09639-25	Sequence 25, Appl
13	1812	98.2	351	7	US-11-088-686-23	Sequence 25, Appl
14	1812	98.2	351	7	US-11-088-686-25	Sequence 25, Appl
15	1811	98.1	351	1	PCT-US05-09639-27	Sequence 27, Appl
16	1811	98.1	351	7	US-11-088-686-27	Sequence 27, Appl
17	1454	78.8	389	1	PCT-US03-10870-559	Sequence 559, App
18	1454	78.8	389	1	PCT-US05-09639-1	Sequence 1, Appl
19	1454	78.8	389	1	PCT-US05-21650-33	Sequence 33, Appl
20	1454	78.8	389	6	US-10-940-774A-6016	Sequence 6016, Ap
21	1454	78.8	389	7	US-11-088-686-1	Sequence 1, Appl
22	1454	78.8	449	6	US-10-990-328A-12349	Sequence 12349, A
23	1454	78.8	449	6	US-10-990-328A-12350	Sequence 12350, A
24	1454	78.8	453	6	US-10-450-763-45646	Sequence 45646, A
25	1454	78.8	453	6	US-10-940-774A-7896	Sequence 7896, Ap
26	1450	78.5	389	1	PCT-US05-09639-3	Sequence 3, Appl
27	1450	78.5	389	7	US-11-088-686-3	Sequence 3, Appl
28	1449	78.5	389	1	PCT-US05-09639-5	Sequence 5, Appl
29	1449	78.5	389	7	US-11-088-686-5	Sequence 5, Appl
30	1446	78.3	389	1	PCT-US05-09639-7	Sequence 7, Appl
31	1445	78.3	389	7	US-11-088-686-7	Sequence 7, Appl
32	1445	78.3	389	1	PCT-US05-09639-9	Sequence 9, Appl
33	1445	78.3	389	1	PCT-US05-09639-11	Sequence 11, Appl
34	1445	78.3	389	7	US-11-088-686-9	Sequence 9, Appl
35	1445	78.3	389	7	US-11-088-686-11	Sequence 11, Appl
36	1444	78.2	389	1	PCT-US05-09639-13	Sequence 13, Appl
37	1444	78.2	389	7	US-11-088-686-13	Sequence 13, Appl
38	1371	74.3	389	8	US-60-669-241-28508	Sequence 28508, A
39	1348	73.0	393	7	US-11-097-143-19182	Sequence 19182, A
40	1274	69.0	319	1	PCT-US05-12947-102	Sequence 102, App
41	1232.5	66.8	404	6	US-10-703-032-123588	Sequence 123588,
42	1156	62.6	330	7	US-11-251-208-421	Sequence 421, App
43	1147.5	62.2	329	7	US-11-251-208-427	Sequence 427, App
44	1128.5	61.1	327	7	US/11/241	GENERAL INFORMATI
45	1125.5	61.0	345	7	US-11-251-208-519	Sequence 519, App

ALIGNMENTS

RESULT 1

PCT-US05-09639-15
; Sequence 15, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004WO1
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; PRIOR FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-09639-15

Alignment Scores:
Pred. No.: 5,1e-178 Length: 351
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.65% Indels: 0
DB: 1 Gaps: 0

US-10-698-228-2 (1-1053) x PCT-US05-09639-15 (1-351)

QY 1 ATGGCGCCACCGGAAAGCCGCGGCTGGATCAGATGAGATCATCTTCA 60
|||||
Db 1 MetGlyaspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSer 20
|||||
QY 61 GACACCAACGAAAGTGAATAAGTCAATGAAGCCACTCTTAAGAAAGTTCTCGC 120

```
Db 21 AspThrAsnGluSerGluLeuIleLysSerAsnGluProLeuLeuArgLysSerSerArg 40
QY 121 CGGTTTGTCACTTCCAAATCCAGTACCTGATATTTGGAAATGATATAACAGGACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCTCTTCGACAGCAGAGAGTTCGACTTATCAAAAGGATCTCCCTCAGTGGAAACA 240
Db 61 AlaserPheThrAlaGluValAspLeuSerLysAspLeuProHisTyrPheLys 80
QY 241 CTTAAAGCAGATGAGAAGTACTCTCTCATCTTAGCCCTTTTTCGACCCAGTGTAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisLeuAlaPhePheAlaAlaSerAsp 100
QY 301 GGAATTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACACTTATCAGAGATCCCAAGAAAGGAAATTTTATTTAATGCAATTTGAA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTAAGAAAAGCAGATTTCCAGAAAAGGAAATTTTATTAATGCAATTTGAA 540
Db 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTTACTTTTGGGAAAGAGTGGTGGCCCTTGTCTGTAGAGGAGTTCCTTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTCTATATTTCTGGCTAAAGAAGAGGTCTTATGCCAGGACTCATTCTTCC 660
Db 201 SerPheAlaAlaIlePheThrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACCTATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTGCTGCTGCTGCTGCT 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAATTAAGCTTCCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTCTCAAA 780
```

RESULT 2

```
PCT-US05-21650-34
; Sequence 34, Application PC/TUS0521650
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: MIGFS AS MODIFIERS OF THE IGF PATHWAY AND METHODS OF USE
```

```
; FILE REFERENCE: EX05-017
; CURRENT APPLICATION NUMBER: PCT/US05/21650
; CURRENT FILING DATE: 2005-06-27
; PRIOR APPLICATION NUMBER: 60/581,689
; PRIOR FILING DATE: 2004-06-21
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-21650-34

Alignment Scores: 51e-178 Length: 351
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservations: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.65% Indels: 0
DB: 1 Gaps: 0

US-10-698-228-2 (1-1053) x PCT-US05-21650-34 (1-351)

QY 1 ATGGCGGACCCCGAAAGCGCGGAGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAACGAAAGTGAATTAAGTCAATGACAGAGCCACTCTCTAAGAAAGATTCTCGC 120
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerSerArg 40
QY 121 CGGTTTGTCACTTCCAAATCCAGTACCTGATATTTGGAAATGATAAAGAGGACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCTCTTCGACAGCAGAGAGTTCGACTTATCAAAAGGATCTCCCTCAGTGGAAACA 240
Db 61 AlaserPheThrAlaGluValAspLeuSerLysAspLeuProHisTyrPheLys 80
QY 241 CTTAAAGCAGATGAGAAGTACTCTCTCATCTTAGCCCTTTTTCGACCCAGTGTAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisLeuAlaPhePheAlaAlaSerAsp 100
QY 301 GGAATTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGGT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACACTTATCAGAGATCCCAAGAAAAGGAAATTTTATTAATGCAATTTGAA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 160
QY 481 ACCATGCCCTATGTAAGAAAAGCAGATTTCCAGAAAAGGAAATTTTATTAATGCAATTTGAA 540
Db 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTTACTTTTGGGAAAGAGTGGTGGCCCTTGTCTGTAGAGGAGTTCCTTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTCTATATTTCTGGCTAAAGAAGAGGTCTTATGCCAGGACTCATTCTTCC 660
Db 201 SerPheAlaAlaIlePheThrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACCTATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTGCTGCTGCTGCTGCT 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAATTAAGCTTCCAGAGAAAGGCTCAGGAGATCATTTGTTGATGCTCTCAAA 780
```

Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluLeuLeValAspAlaValLys 260
QY 781 ATTGACGAGGAGTTTAAACAGAGCTTGCAGAGTGGCTCATTTGGAATGATTCATT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAACACGATGAGTTGTAGCTGACAGATTAATCTTGTGGAATTTGGATTCCTCA 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTCAGCAGAAATCCTTTTGATTTTATGGAACAAATTTCTTTAGAGAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAAAATTTCTTTCAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGAGTATGCGAGAAACC 1020
Db 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGTTCTTACCTTGGATGCGAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 3
US-10-990-328A-7547
; Sequence 7547, Application US/10990328A
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328A
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7547
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-990-328A-7547

Alignment Scores:
Pred. No.: 5,1e-178 Length: 351
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.65% Indels: 0
DB: 6 Gaps: 0

US-10-698-228-2 (1-1053) x US-10-990-328A-7547 (1-351)

QY 1 ATGGCGACCCCGAAGCGCGGCGGCTGGATCAGATGAGAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAACGAGTGAATTAAGTCAATGAGAGCCACTCTTAAGAAGAGTTCTCC 120
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
QY 121 CGGTTTGTTCATCTTCCATCAGTACCTGATATTGGAATATGATAAAGCCAGAC 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCCTTCTGACAGCAGAGAGGTTCGATTCATCAAGGATCTCCCTCAGTGAACAG 240
Db 61 AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
QY 241 CTTAAGCAGATGAGAGTACTTCTCTCATCATCTTACGCTTTTTCGACCCAGTGT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAsp 100
QY 301 GGAATTTGAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGCTCCAGGCT 360

Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATGCTTTCAAATTCATCATCAGAGATGTTCACTCAGAGATGTACAGTTG 420
Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAAGGAATTTTATTAATCAATGAA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTAAGAAAAAGCAGATTGGGCTTGGCGATGATGATGATGATGAA 540
Db 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAGAGTGGCTTCTCTCTAGAGAGGATTTCTTCTTCAGCA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTCTGCTATATTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACTTTTCC 660
Db 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATCAACTCATCAGCAGAGATGAAGACTTCACTGTGCTTGTGCTGCTGATGTTCCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAATAAGCTTTCAGAAAGAGGTGAGGAGATCATTTGATGATGCTGCTCAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
QY 781 ATTGAGCAGGAGTTTAAACAGAGCTTGCAGTGGCCCTCATTTGGAATGATTCATT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAACACGATGAGTTGTAGCTGACAGATTAATCTTGTGGAATTTGATTCCTCA 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGTTTTTTCAGCAGAAATCCTTTTGATTTTATGGAACAAATTTCTTTAGAGAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAAAATTTCTTTCAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGAGTATGCGAGAAACC 1020
Db 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGCTTCTTACCTTGGATGCGAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 4
US-11-088-686-15
; Sequence 15, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-088-686-15

Alignment Scores:
Pred. No.: 5,1e-178 Length: 351
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.65% Indels: 0
DB: 7 Gaps: 0

US-10-698-228-2 (1-1053) x US-11-088-686-15 (1-351)

```
QY 1 ATGGCGACCCGGAAGCCGGAAGCGCGCGCTGGATCAGAGATCATCTTCA 60
DB 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAACGAAAGTAAATGAATCAATGAAGAGCCACTCTTAAGAAAGTCTCGC 120
DB 21 AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerArg 40
QY 121 CGGTTTGTCACTTCCCAATCCAGTACCTGATATTTGGAATGATATTAACAGCAG 180
DB 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrPheMetTyrLysGlnAla 60
QY 181 GCTTCTTCTGACAGCAGAGAGTCCGACTTATCAAGGATCTCCCTCACTGGAAAC 240
DB 61 AlaSerPheTyrThrAlaGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
QY 241 CTTAAAGCAGATGAAAGTACTTCTCACTTCTCACTTCTTCTTCTTCTTCTTCT 300
DB 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp 100
QY 301 CGAATGTAAATGAATTTGGTGGAGCGCTTGTAGTCAGGAGTCCAGGTCAGAGCT 360
DB 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnProGluAla 120
QY 361 CGCTGTTCTATGCTTTCCTCAATCTCTCAGAGATGTTCTCACTCAGAGATGTAC 420
DB 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSer 140
QY 421 CTGATPAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATG 480
DB 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIle 160
QY 481 ACCATGCCCTATGTTAAGAAAGAGATTTGGCTTGTAGAGAGTCTTCTTCTCAG 540
DB 161 ThrMetProTyrValLysLysLysAlaAspTyrAlaLeuArgTyrIleAlaAspArg 180
QY 541 TCTACTTTGGGGAAGAGTGTGGCTTGTCTGTGTAGAGAGTCTTCTTCTCAGCA 600
DB 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSer 200
QY 601 TCTTTTGTCTATATTCCTGCTTAAAGAGAGAGTCTTATCCAGGACTCACTTTTCC 660
DB 201 SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPhe 220
QY 661 ATGAACTCATCAGCAGATGAGGACTTCACTGTGACTTGTCTTGTCTGCTGATGTC 720
DB 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPhe 240
QY 721 TACTTAGTAAATAAGCTTCAAGAAAGAGGTCCAGGAGATCATCTTCTGCTGTCA 780
DB 241 TyrLeuValAsnLysProSerGluArgValArgGluIleValAspAlaValLys 260
QY 781 ATGTAGCAGGAGTCTTTTAAAGAGAGCTTGCAGAGTGTGCTCATTTGGAATGAT 840
DB 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCys 280
QY 841 TTGATGAAACAGTACATTGAGTTGTAGCTGACAGATTACTTGTGGAACCTTGGAT 900
DB 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPhe 300
QY 901 AAGGTTTTTCGCGCAGAAATCTTTGATTTATGGAAGAAATTTCTTTAGAGGAAA 960
DB 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluLys 320
QY 961 ACAAAATTTCTTTGAGAAACAGTTTTCAGAGTATCAGCGTTTTCAGTATGCGAGA 1020
DB 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGlu 340
```

```
QY 1021 ACAGATAACGCTCTTCCACCTTGGATGCAGATTTT 1053
DB 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
```

RESULT 5

PCT-US05-09639-17
; Sequence 17, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004WOI
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-09639-17

Alignment Scores:
Pred. No.: 1,31e-177 Length: 351
Score: 1817.00 Matches: 350
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.72% Mismatches: 0
Query Match: 98.43% Indels: 0
DB: 1 Gaps: 0

US-10-698-228-2 (1-1053) x PCT-US05-09639-17 (1-351)

```
QY 1 ATGGCGACCCGGAAGCCGGAAGCGCGCGCTGGATCAGAGATCATCTTCA 60
DB 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGTCTCGC 120
DB 21 AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerSerArg 40
QY 121 CGGTTTGTCTATCTTCCCAATCCAGTACCTCTGATATTTGGAATGATATTAACAGCAG 180
DB 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrPheMetTyrLysGlnAla 60
QY 181 GCTTCTTCTGACAGCAGAGAGTCTGATTAACAAGGATCTCCCTCACTGGAAAC 240
DB 61 AlaSerPheTyrThrAlaGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
QY 241 CTTAAAGCAGATGAGAGTACTTCTCACTCTCACTCTTAGCCTTTTTCGAGCCAGTCA 300
DB 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp 100
QY 301 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTCCAGGTCAGAGCT 360
DB 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnProGluAla 120
QY 361 CGCTGTTTCTATGCTTTCCTCAATTTCTCATCGAATGTTCCTCAGAGATGTACAGTTG 420
DB 121 ArgCysPhePheGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSer 140
QY 421 CTGATPAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATG 480
DB 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIle 160
QY 481 ACCATGCCCTATGTTAAGAAAGAGATTTGGCTTGTAGAGAGTCTTCTTCTCAG 540
DB 161 ThrMetProTyrValLysLysLysAlaAspTyrAlaLeuArgTyrIleAlaAspArg 180
QY 541 TCTACTTTTGGGGAAGAGTGTGGCTTGTCTGTGTAGAGAGTCTTCTTCTCAGCA 600
DB 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSer 200
QY 601 TCTTTTGTCTATATTCCTGCTTAAAGAGAGAGTCTTATCCAGGACTCACTTTTCC 660
DB 201 SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPhe 220
QY 661 ATGAACTCATCAGCAGATGAGGACTTCACTGTGACTTGTCTTGTCTGCTGATGTC 720
DB 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPhe 240
QY 721 TACTTAGTAAATAAGCTTCAAGAAAGAGGTCCAGGAGATCATCTTCTGCTGTCA 780
DB 241 TyrLeuValAsnLysProSerGluArgValArgGluIleValAspAlaValLys 260
QY 781 ATGTAGCAGGAGTCTTTTAAAGAGAGCTTGCAGAGTGTGCTCATTTGGAATGAT 840
DB 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCys 280
QY 841 TTGATGAAACAGTACATTGAGTTGTAGCTGACAGATTACTTGTGGAACCTTGGAT 900
DB 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPhe 300
QY 901 AAGGTTTTTCGCGCAGAAATCTTTGATTTATGGAAGAAATTTCTTTAGAGGAAA 960
DB 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluLys 320
QY 961 ACAAAATTTCTTTGAGAAACAGTTTTCAGAGTATCAGCGTTTTCAGTATGCGAGA 1020
DB 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGlu 340
```

```
QY 601 TCTTTGCTGCTATATCTTGGCTAAAGAGAGAGAGTCTTATGCGAGAGTCACTTTTCC 660
Db 201 SerPheAlaAlaIlePheTrpLeuLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACATCATCAGCGAGATGAGAGCTTCACTGTGATCTTGGCTTGGCTGATGTTCCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAATAAGCCTTCAGAGAAAGGTCAGGAGATCATCTTGTGATGCTGTCMAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleValAspAlaValLys 260
QY 781 ATTGAGCAGAGTCTTTTAAACAGAGCCTTGGCAGTGGCTCATTTGGAATGAATTCATT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAACACATGATGAGTTGCTGAGTGTGAGATTAATCTTGTGAACTTGGATTTCTCA 900
Db 281 LeuMetLysGlnTyrlleGluPheValAlaAspArgLeuLeuGluLeuGlyPheSer 300
QY 901 AAGCTTTTTCAGCGAGAAATCCTTTTGAATTTATGGAACATTTCTTTAGAGAAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAATTTCTTGGAGAAACGAGTTTCAGAGTATCAGCGTTTGGCAGTATGCGAGAAACC 1020
Db 321 ThrAsnPheGluLysArgValSerGluTyrlleGluArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
```

RESULT 6

US-11-088-686-17

; Sequence 17, Application US/11088686

; GENERAL INFORMATION:

; APPLICANT: Yen, Yun

; TITLE OF INVENTION: DRUG SCREENING

; FILE REFERENCE: 14037-004001

; CURRENT APPLICATION NUMBER: US/11/088,686

; CURRENT FILING DATE: 2005-03-23

; PRIOR APPLICATION NUMBER: US 60/556,836

; PRIOR FILING DATE: 2004-03-25

; NUMBER OF SEQ ID NOS: 59

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 17

; LENGTH: 351

; TYPE: PRT

; ORGANISM: Homo sapiens

US-11-088-686-17

Alignment Scores:

Pred. No.:	1,316-177	Length:	351
Score:	1817.00	Matches:	350
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.72%	Mismatches:	0
Query Match:	98.43%	Indels:	0
DB:	7	Gaps:	0

US-10-698-228-2 (1-1053) x US-11-088-686-17 (1-351)

```
QY 1 ATGGCGAGCCGGAAGCGCGGAGCGGCTGATCAGATGAGAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGluAspGluArgSerSer 20
QY 61 GACACCAAGCAAGTGAATAAGTCAATGAAGCCACTCTTAAGAAGAGTTCGC 120
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerArg 40
QY 121 CGGTTTGTCTATCTTCCATCAGTACCTGCTGATATTTGGAAATGATATAACAGGCACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrlleProAspIleTrpLysGlnAlaGln 60
```

RESULT 7

PCT-US05-09639-19

; Sequence 19, Application PC/TUS0509639

; GENERAL INFORMATION:

; APPLICANT: Yen, Yun

; TITLE OF INVENTION: DRUG SCREENING

; FILE REFERENCE: 14037-004001

; CURRENT APPLICATION NUMBER: PCT/US05/09639

; CURRENT FILING DATE: 2005-03-23

; PRIOR APPLICATION NUMBER: US 60/556,836

; PRIOR FILING DATE: 2004-03-25

; NUMBER OF SEQ ID NOS: 59

```
QY 181 GCTTCTCTTCGAGCAGCAGAGAGGTGAGTCTTATCAAAAGGATCTCCCTCACTGGAACAAG 240
Db 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAAGCAGATGAGAAAGTACTTCACTCTCAATCTTAGCCCTTTTGGACCCAGTGAT 300
Db 81 LeuLysAlaAspGluLysTyrlleSerHisIleLeuAlaPhePheAlaAspAsp 100
QY 301 GGAATTGTAATGAAATTTGGTGGAGCCTTTAGTCAGAGGTGCGAGTTCCAGAGGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATGCTTCAAAATCTCATCAGAAATGTTCACCTCAGAGATGTACAGATTG 420
Db 121 ArgCysPhePheGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrlleSerLeu 140
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAAATGCAATTGAA 480
Db 141 LeuIleAspThrTyrlleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATGGCCCTTGGATCGATGATGAGATGAGAAA 540
Db 161 ThrMetProTyrlleValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCCTTTGCTCTGTAGAGAGGATTTTCTTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTCTCTATATTCTGGCTAAAGAGAGAGTCTTATGCCAGGACTCACATTTTCC 660
Db 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACATCATCAGCAGATGAGAGCTTCACTGTGATCTTGTGCTGCTGATGTTCCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAATAAGCCTTCAGAGAAAGGTGAGGAGATCATCTTGTGCTGCTCTCAAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleValAspAlaValLys 260
QY 781 ATTGAGCAGGAGTCTTTTAAACAGAGCCTTGCAGTGGCCTCATTTGGAATGAATTCATT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAACACATGATGAGTTGTTGATGCTGACAGATTAATCTTGTGAACTTGGATTC 900
Db 281 LeuMetLysGlnTyrlleGluPheValAlaAspArgLeuLeuGluLeuGlyPheSer 300
QY 901 AAGCTTTTTCAGCGAGAAATCCTTTTGAATTTTATGGAACATTTCTTTAGAGAAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAATTTCTTGGAGAAACGAGTTTCAGAGTATCAGCGTTTGGCAGTATGCGAGAAACC 1020
Db 321 ThrAsnPheGluLysArgValSerGluTyrlleGluArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-09639-19

Alignment Scores:
Pred. No.: 1 666-177 Length: 351
Score: 1816.00 Matches: 350
Percent Similarity: 100.00% Conservatives: 1
Best Local Similarity: 99.72% Mismatches: 0
Query Match: 98.37% Indels: 0
DB: 1 Gaps: 0

US-10-698-228-2 (1-1053) x PCT-US05-09639-19 (1-351)

QY 1 ATGGCGGACCCGGAAGCGCGGCTGGATCAGATCAGAGATCATCTTCA 60
DB 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAAGCAAGTGAATAAGTCAATGAAGACCCACTCTAAGAAAGTCTCGC 120
DB 21 AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerArg 40
QY 121 CGTTTGTTCATCTTCCCAATCCAGTACCCTGATATTTGGAATATGTAACAGGCACAG 180
DB 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCCTTCTGACAGCAGAGAGGTGCTGATTCATCAAGGATCTCCCTCAGTCAAGCAAG 240
DB 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAGCAGATGAGAGTACTCTCTCATCTCAGATCTTACCTTTTTCAGCCAGTGTAT 300
DB 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAlaSerAsp 100
QY 301 GGAATTTGAATAAGTAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGTTCAGAGGCT 360
DB 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATGGCTTTTCAAAATTCATCGAGAATGTTTCACCTCAGAGATGTACAGTTTG 420
DB 121 ArgCysPheTrpGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
QY 421 CTGATAGACACTATCATCAGATGCCAAGAAAGGAAATTTTATTTATGCAATTGAA 480
DB 141 LeuIleAspThrTyrIleArgAspProLysArgGluPheLeuPheAlaAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAGAAAGACAGATTCGGCCCTTCGATGGATAGCAGTAAAAA 540
DB 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAGAGTCTTCTTCAGGA 600
DB 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTGCTATATTCGGCTTAAAGAGAGAGGTCTTATGCGCAGACTCATCTTTTTC 660
DB 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 ATGACTCATCAGCAGATGAAGCACTTCTGCTGACTTGTGCTTGTGCTGCTGCTGCTTCCAA 720
DB 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAAATGAAGCTTCAGAAGAAAGGTTCAGGAGATCATTTGTGTGCTGTGCAAA 780
DB 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
QY 781 ATTGACGAGAGTATTTTAAACAGAGCCTTTCAGAGTTCGCTCATTTGGAATGATTCATT 840
DB 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
```

```
QY 841 TTGATGAAACAGTACATTTAGTTTGTAGCTGACAGATTACTTGTGGAATTTGGATTCTCA 900
DB 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGTTTTTCAGCAGAAAAATCCTTTTGTATTTATGGAACAAATTTCTTTAGAGGAAAA 960
DB 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAAAATTTCTTTCAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAAACC 1020
DB 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGCTCTTCACCTTTGGATGCGAGATTTT 1053
DB 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 8
US-11-088-686-19
; Sequence 19, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-088-686-19

Alignment Scores:
Pred. No.: 1 666-177 Length: 351
Score: 1816.00 Matches: 350
Percent Similarity: 100.00% Conservatives: 1
Best Local Similarity: 99.72% Mismatches: 0
Query Match: 98.37% Indels: 0
DB: 1 Gaps: 0

US-10-698-228-2 (1-1053) x US-11-088-686-19 (1-351)

QY 1 ATGGCGGACCCGGAAGCGCGGCTGGATCAGATCAGAGATCATCTTCA 60
DB 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAAGCAAGTGAATAAGTCAATGAAGACCCACTCTAAGAAAGTCTCGC 120
DB 21 AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerArg 40
QY 121 CGTTTGTTCATCTTCCCAATCCAGTACCCTGATATTTGGAATATGTAACAGGCACAG 180
DB 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCCTTCTGACAGCAGAGGTGCTGATTCATCAAGGATCTCCCTCAGTCAAGCAAG 240
DB 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAGCAGATGAGAGTACTCTCTCATCTCAGATCTTACCTTTTTCAGCCAGTGTAT 300
DB 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAlaSerAsp 100
QY 301 GGAATTTGAATAAGTAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGTTCAGAGGCT 360
DB 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATGGCTTTTCAAAATTCATCGAGAATGTTTCACCTCAGAGATGTACAGTTTG 420
DB 121 ArgCysPheTrpGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
```



```
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAATTTTATTTAATGCAATGAA 480
Db 141 LeuileAspThrTyrileArgAspProLysLysArgGluPheLeuPheAsnAlaileGlu 160
QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGCCCTTCGGATGATACAGATAGAAA 540
Db 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpilealaAspArgLys 180
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGTGTAGAGGAGTCTTCTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTGCTATATCTTGGCTAAAGAGAGAGTCTTATGCCAGTCTTATGCTTTTCC 660
Db 201 SerPheAlaAlailePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACATCAGCAGAGATGAAGACTTCTACTGTGACTTTGCTTGGCTGTGATGTTCCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAATAAGCTTCAGAAAGAGGTCAGGAGATCATTTGATGCTGTCAAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluileileValAspAlaValLys 260
QY 781 ATTGAGCAGGAGTTTAAACAGAGCTTTCAGAGTTCAGTGTGACTTTGCTTGGCTGTGATGCTCAAT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuileGlyMetAsnCysIle 280
QY 841 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTAATTTGTAATGAACTTGGATTTCTCA 900
Db 281 LeuMetLysGlnTyrileGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTCAGGCAGAAAATCCTTTTGGTATGTAAGAGGATCATTTGATGCTGTCAAA 960
Db 301 LysValPheGlnAlailePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 320
QY 1021 ACAGATAACGTCCTTACCTTGGATGAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
```

RESULT 9

```
PCT-US05-09639-21
; Sequence 21, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen. Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; PRIOR FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-09639-21
```

```
Alignment Scores:
Pred. No.: 3,38e-177 Length: 351
Score: 1813.00 Matches: 349
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.43% Mismatches: 0
Query Match: 98.21% Indels: 0
DB: 1 Gaps: 0
```

US-10-698-228-2 (1-1053) x PCT-US05-09639-21 (1-351)

```
QY 1 ATGGCGCACCGCGAAAGCGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGluArgSerSerSer 20
QY 61 GACACCAACGAAAGTGAATTAAGTCAATGAGAGCCACTCCTTAAGAAAGATTCGCG 120
Db 21 AspThrAsnGluSerGluileLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
QY 121 CGGTTTGTCTATCTTCCAAATCCAGTACCTCGATATTTGGAAAAATGTAACACGCACAG 180
Db 41 ArgPheValilePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCTCTTCGACAGCAGAGAGGTCAGCTTATCAAGAGTCTCCCTCAGCTGCAACAAG 240
Db 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAAGCAGATCAGAAAGTACTTCTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTGAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisileLeuAlaPhePheAlaAspAsp 100
QY 301 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGGAGTGCAGGTTCCAGAGCT 360
Db 101 GlyileValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATCGCTTTTCAAAATCTCATCGAATGTTCTCAGAGATGTACAGTTTG 420
Db 121 ArgCysPhePheGlyPheGlnIleLeuileGluAsnValHisSerGluMetPheSerLeu 140
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAATTTTATTTAATGCAATGAA 480
Db 141 LeuileAspThrTyrileArgAspProLysLysArgGluPheLeuPheAsnAlaileGlu 160
QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGCCCTTCGGATGATACAGATAGAAA 540
Db 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpilealaAspArgLys 180
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGTCTAGAGGATTTCTTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGGCTTATGCGAGGACTCACATTTTTC 660
Db 201 SerPheAlaAlailePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACATCAGCAGAGATGAAGACTTCTACTGTGACTTTGCTTGGCTGTGATTTCCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAATAAGCTTCAGAAAGAGGTCAGGAGATCATTTGATGCTGTCAAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluileileValAspAlaValLys 260
QY 781 ATTGAGCAGGAGTTTAAACAGAGCTTTCAGAGTTCAGTGTGACTTTGTAATGAACTTGA 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuileGlyMetAsnCysIle 280
QY 841 TTGATGAACAGTACATGAGTTTGTAGCTGACAGATTAATTTGTAATGAACTTGGATTTCTCA 900
Db 281 LeuMetLysGlnTyrileGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTCAGGCAGAAAATCCTTTTGGTATGTAAGAGGATCATTTTGTAGAGGAAA 960
Db 301 LysValPheGlnAlailePheTrpLeuLysLysArgPhePheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAAAATTTCTTTGAGAAACAGTTTTCAGAGTATCAGCGTTTTCAGTATGCGAGAAACC 1020
Db 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGTCCTTACCTTGGATGAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
```

RESULT 10

```
US-11-088-686-21
; Sequence 21, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-088-686-21

Alignment Scores:
Pred. No.: 3,38e-177 Length: 351
Score: 1813.00 Matches: 349
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.43% Mismatches: 0
Query Match: 98.21% Indels: 0
DB: 7 Gaps: 0

US-10-698-228-2 (1-1053) x US-11-088-686-21 (1-351)

QY 1 ATGGGCGACCGGAAGCGCGGAGCGCGGCTGGATCAGATCAGAGATCATCTTCA 60
DB 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAACGAAGTCAATTAAGTCAATGAAGCCCTCTCAAGAAAGAGTCTCGC 120
DB 21 AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerSer 40
QY 121 CGTTTGTCTCATCTTCCATCCAGTACCTCTGATATTGGAATATGTAACAGGACAG 180
DB 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAla 60
QY 181 GCTTCTCTTGGACAGAGAGGTGCGATATCAAGAGTCTCCCTCACTCGAACAAG 240
DB 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAGCAGATCAGAGTACTTCTCATCTCTACATCTTACCTTTTTCGAGCCAGT 300
DB 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp 100
QY 301 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 360
DB 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATGGCTTTCAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTG 420
DB 121 ArgCysPhePheGlyPheGlnIleLeuIleGluAsnValHisSerGluMetPheSerLeu 140
QY 421 CTCATAGACATCTACATCAGAGATCCCAAGAAAGGAAATTTTATTATGCAATGAA 480
DB 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATCTTAAGAAAAAGCAGATTTGGCGCTTGGCATGGATAGCAGATAGAAA 540
DB 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAGAGGATTTTCTTCTCAGGA 600
DB 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGluValPhePheSerGly 200
QY 601 TCTTTTGTCTGCTATATTCTGGCTAAAGACAGAGGTCTTATGCCAGGACCTCACTTTTCC 660
DB 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTTCTGCTGCTGATGTTCCAA 720

US-11-088-686-21
; Sequence 21, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-088-686-21

Alignment Scores:
Pred. No.: 3,38e-177 Length: 351
Score: 1813.00 Matches: 349
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.43% Mismatches: 0
Query Match: 98.21% Indels: 0
DB: 7 Gaps: 0

US-10-698-228-2 (1-1053) x PCT-US05-09639-23 (1-351)

QY 1 ATGGGCGACCGGAAGCGCGGAGCGCGGCTGGATCAGATCAGAGATCATCTTCA 60
DB 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAACGAAGTCAATTAAGTCAATGAAGCCCTCTCAAGAAAGAGTCTCGC 120
DB 21 AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerSer 40
QY 121 CGTTTGTCTCATCTTCCATCCAGTACCTCTGATATTGGAATATGTAACAGGACAG 180
DB 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAla 60
QY 181 GCTTCTCTTGGACAGAGGTGCGATATCAAGAGTCTCCCTCACTCGAACAAG 240
DB 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAGCAGATCAGAGTACTTCTCATCTCTACATCTTACCTTTTTCGAGCCAGT 300
DB 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp 100
QY 301 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 360
DB 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATGGCTTTCAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTG 420
DB 121 ArgCysPhePheGlyPheGlnIleLeuIleGluAsnValHisSerGluMetPheSerLeu 140
QY 421 CTCATAGACATCTACATCAGAGATCCCAAGAAAGGAAATTTTATTATGCAATGAA 480
DB 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATCTTAAGAAAAAGCAGATTTGGCGCTTGGCATGGATAGCAGATAGAAA 540
DB 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAGAGGATTTTCTTCTCAGGA 600
DB 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGluValPhePheSerGly 200
QY 601 TCTTTTGTCTGCTATATTCTGGCTAAAGACAGAGGTCTTATGCCAGGACCTCACTTTTCC 660
DB 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTTCTGCTGCTGATGTTCCAA 720

PCT-US05-09639-23
; Sequence 23, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-09639-23

Alignment Scores:
Pred. No.: 4,29e-177 Length: 351
Score: 1812.00 Matches: 349
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.43% Mismatches: 0
Query Match: 98.16% Indels: 0
DB: 1 Gaps: 0

US-10-698-228-2 (1-1053) x PCT-US05-09639-23 (1-351)

QY 1 ATGGGCGACCGGAAGCGCGGAGCGCGGCTGGATCAGATCAGAGATCATCTTCA 60
DB 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAACGAAGTCAATTAAGTCAATGAAGCCCTCTCAAGAAAGAGTCTCGC 120
DB 21 AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerSer 40
QY 121 CGTTTGTCTCATCTTCCATCCAGTACCTCTGATATTGGAATATGTAACAGGACAG 180
DB 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAla 60
QY 181 GCTTCTCTTGGACAGAGGTGCGATATCAAGAGTCTCCCTCACTCGAACAAG 240
DB 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
QY 241 CTTAAGCAGATCAGAGTACTTCTCATCTCTACATCTTACCTTTTTCGAGCCAGT 300
DB 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp 100
QY 301 GGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 360
DB 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATGGCTTTCAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTG 420
DB 121 ArgCysPhePheGlyPheGlnIleLeuIleGluAsnValHisSerGluMetPheSerLeu 140
QY 421 CTCATAGACATCTACATCAGAGATCCCAAGAAAGGAAATTTTATTATGCAATGAA 480
DB 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATCTTAAGAAAAAGCAGATTTGGCGCTTGGCATGGATAGCAGATAGAAA 540
DB 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAGAGGATTTTCTTCTCAGGA 600
DB 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGluValPhePheSerGly 200
QY 601 TCTTTTGTCTGCTATATTCTGGCTAAAGACAGAGGTCTTATGCCAGGACCTCACTTTTCC 660
DB 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTTCTGCTGCTGATGTTCCAA 720
```

```
Db      81  LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY      301  CGAATTGTAATAATCGAAATTTGGTGGAGCGCTTTAGTCAGGAGTCCAGAGGCT 360
Db      101  GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY      361  CGCTGTTCTATGGCTTCCAAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db      121  ArgCysPhePheGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTrpSerLeu 140
QY      421  CTGATAGACATCATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 480
Db      141  LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY      481  ACCATGCCCTATGTTAAGAAAGACAGATTCGGCTTCGGATGGATACAGATAGAAA 540
Db      161  ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY      541  TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAAGAGTTCCTTCTCAGGA 600
Db      181  SerThrPheGlyGluArgValValAlaPheAlaIleValGluGlyValPhePheSerGly 200
QY      601  TCTTTTGTGCTATATTTCTGGCTTAAGAAAGAGAGTCTTTATGCCAGACTCACTTTTCC 660
Db      201  SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY      661  AATGAACATCATCAGAGATGAGGACTTCCTACTGTGACTTTGCTGGCTGTATGTTCCAA 720
Db      221  AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY      721  TACTTAGTAAATAAGCTTCAGAGAAAGGCTTCAGGAGATCATCTGTCGTGCAAA 780
Db      241  TyrLeuValAsnLysProSerGluArgValArgValArgLysIleValAspAlaValLys 260
QY      781  ATTGACGAGGAGTTCCTTAACAGAAAGCTTCGCTTGAAGAGTTCATTCGGAATGCAAT 840
Db      261  IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY      841  TTGATGAACAGTACATTGAGTTGTAGCTGACAGATTACTTGTGGAACCTTGATTTCTCA 900
Db      281  LeuMetLysGlnTyrIleGluPheValAlaAspPheMetGluAsnIleSerLeuGluGlyLys 300
QY      901  ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTATGCGCAGAAACC 1020
Db      321  ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY      1021  ACAGATAACGCTTTCACCTTGGATCGAGATTTT 1053
Db      341  ThrAspAsnValPheThrLeuAlaAspPhe 351
```

RESULT 12

```
PCT-US05-09639-25
; Sequence 25, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-09639-25
```

```
Alignment Scores:
Pred. No.: 4,29e-177 Length: 351
Score: 1812.00 Matches: 349
Percent Similarity: 100.00% Conservations: 2
Best Local Similarity: 99.43% Mismatches: 0
Query Match: 98.16% Indels: 0
DB: 1 Gaps: 0
US-10-698-228-2 (1-1053) x PCT-US05-09639-25 (1-351)
```

```
QY      1  ATGGGCGACCCGAAAGCCGCGAGCGCGGTGGATCAGGATGAGAGATCATCTTCA 60
Db      1  MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY      61  GACACCAACGAAAGTGAATAAGTCAATCAAGAGCCACTCTTAAGAAAGATTCTCGC 120
Db      21  AspThrAsnGluSerGluIleLysSerAsnGluProLeuLeuArgLysSerSerArg 40
QY      121  CGGTTTGTCTCATCTTCCAAATCCAGTACCCTGATATTTGGAAATGTATAAACAGC 180
Db      41  ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAla 60
QY      181  GCTTCTCTTCTGGACAGCAGAGGCTCGACTTATCAAGAGATCTCCCTCACTGGAA 240
Db      61  AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsn 80
QY      241  CTTAAAGCAGATGAGAAAGTACTTCTCATCTCACATCTTAGCCTTTTTCAGCCAG 300
Db      81  LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSer 100
QY      301  GGAATTGTAATAATTTGGTGGAGCGCTTTAGTCAGGAGTCCAGAGGCT 360
Db      101  GlyIleValAsnGluAsnLeuValArgPheSerGlnGluValGlnValProGluAla 120
QY      361  CGCTGTTTCTATGGCTTTCAAATTCATCAGAAATGTTCTACTCAGAGATGTACAG 420
Db      121  ArgCysPheTrpGlyPheGlnIleLeuIleGluAsnValHisSerGluMetPheSer 140
QY      421  CTGATAGACATTCATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 480
Db      141  LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIle 160
QY      481  ACCATGCCCTATGTTAAGAAAGACAGATTCGGCTTCGGATGGATGAGATAGAAA 540
Db      161  ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArg 180
QY      541  TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAAGAGTTCCTTCTCAGGA 600
Db      181  SerThrPheGlyGluArgValValAlaPheAlaIleValGluGlyValPhePheSer 200
QY      601  TCTTTTGTGCTATATTTCTGGCTTAAGAAAGAGGCTTCATGCCAGGACTCACAT 660
Db      201  SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPhe 220
QY      661  AATGAACATCATCAGAGATGAGGACTTCCTACTGTGACTTTGCTGGCTGTATGTT 720
Db      221  AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPhe 240
QY      721  TACTTAGTAAATAAGCTTCAGAGAAAGGCTTCAGGAGATCATCTGTTGATGCTG 780
Db      241  TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaVal 260
QY      781  ATTGACGAGGAGTTCCTTAACAGAAAGCTTCGCTTGAAGAGTTCATTCGGAATG 840
Db      261  IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCys 280
QY      841  TTGATGAACAGTACATTGAGTTGTAGCTGACAGATTACTTGTGGAACCTTGATTT 900
Db      281  LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPhe 300
QY      901  AAGTTTTCAGGACAGAAATCCTTTGATTTTATGGAACATTTCTTTAGAGGAAA 960
```

```
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyIlys 320
QY 961 ACAAAATTTCTTGGAGAAAGAGTTTCAGAGTATCAGCGTTTGGAGTTTATGCGAGAAACC 1020
Db 321 ThrAsnPheGluGlyArgValSerGluGlyGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGTTTCCACCTTGGATCGAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 13
US-11-088-686-23
; Sequence 23, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; PRIOR FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-088-686-23

Alignment Scores:
Pred. No.: 4,29e-177 Length: 351
Score: 1812.00 Matches: 349
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.43% Mismatches: 0
Query Match: 98.16% Indels: 0
DB: 7 Gaps: 0

US-10-698-228-2 (1-1053) x US-11-088-686-23 (1-351)

QY 1 ATGGCGAGCCCGAAGAGCGGAGCGGCTGGATCAGATGAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAAGAAAGTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
QY 121 CGTTTGTGTCATCTTCCAAATCCAGTACCTGTATTTGGAAAATGTATAACAGGCACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCCTTCTGACAGCAGAGAGTGCAGTATCAAGGATCTCCCTCAGTGGAAACAG 240
Db 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
QY 241 CTTAAGCAGATGAGAAGTACTCTCTCATCTTAGCTTTTTCAGCAGTGTAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY 301 GGAATGTAAATGAAAATTTGTGGAGCGCTTTAGTCAGAGGTGCGAGTTCCAGAGGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTTTTCTATGCGTTTCAAAATCTCATCGAGAATGTTTCATCAGAGATGACAGTTTG 420
Db 121 ArgCysPhePheGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTrpSerLeu 140
QY 421 CTGATAGACACTTACATCAGATCCCAGAAAGGAATTTTATTATTCGAATTGAA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAAGAAAAAGCAGATTTGGGCTTCCGATGGATAGCAGTAAAAA 540
Db 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGAGAAAGAGTGGGCTTTGCTGCTGTAGAGAGGAGTTTCTTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACATTTTTC 660
Db 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAATCATCAGCAGAGATGAAGACTTCACCTGTGACTTTGCTTCCCTGATCTTCCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAAATAAGCCTTCAGAAAGAAAGGTCAGGAGATCATTTGTGTGCTGTCAAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleLeuValAspAlaValLys 260
QY 781 ATTGACGAGGAGTTTAAACAGAACGCTTGCAGTTGGCTCATTTGGAATGAATTCGATT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAAAACAGTACATTCAGTTGTTGCTGACGACAGATTAATCTGTGGAACCTTGA 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTCAGGCAGAAAATCCTTTTGTATTTTATGGAACACATTTCTTTAGAAGGAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAAAATTTCTTTCAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGCAGAAACC 1020
Db 321 ThrAsnPheGluGlyLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGTTTCCACCTTGGATCGAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351

RESULT 14
US-11-088-686-25
; Sequence 25, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; PRIOR FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-088-686-25

Alignment Scores:
Pred. No.: 4,29e-177 Length: 351
Score: 1812.00 Matches: 349
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.43% Mismatches: 0
Query Match: 98.16% Indels: 0
DB: 7 Gaps: 0

US-10-698-228-2 (1-1053) x US-11-088-686-25 (1-351)

QY 1 ATGGCGAGCCCGAAGAGCGGAGCGGCTGGATCAGATGAGATCATCTTCA 60
Db 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
QY 61 GACACCAAGAAAGTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
QY 121 CGTTTGTGTCATCTTCCAAATCCAGTACCTGTATTTGGAAAATGTATAACAGGCACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCCTTCTGACAGCAGAGAGTGCAGTATCAAGGATCTCCCTCAGTGGAAACAG 240
Db 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
QY 241 CTTAAGCAGATGAGAAGTACTCTCTCATCTTAGCTTTTTCAGCAGTGTAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY 301 GGAATGTAAATGAAAATTTGTGGAGCGCTTTAGTCAGAGGTGCGAGTTCCAGAGGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTTTTCTATGCGTTTCAAAATCTCATCGAGAATGTTTCATCAGAGATGACAGTTTG 420
Db 121 ArgCysPhePheGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTrpSerLeu 140
QY 421 CTGATAGACACTTACATCAGATCCCAGAAAGGAATTTTATTATTCGAATTGAA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTAAAGAAAAAGCAGATTTGGGCTTCCGATGGATAGCAGTAAAAA 540
Db 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
QY 541 TCTACTTTTGGGAGAAAGAGTGGGCTTTGCTGCTGTAGAGAGGAGTTTCTTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACATTTTTC 660
Db 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGAATCATCAGCAGAGATGAAGACTTCACCTGTGACTTTGCTTCCCTGATCTTCCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAAATAAGCCTTCAGAAAGAAAGGTCAGGAGATCATTTGTGTGCTGTCAAA 780
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleLeuValAspAlaValLys 260
QY 781 ATTGACGAGGAGTTTAAACAGAACGCTTGCAGTTGGCTCATTTGGAATGAATTCGATT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAAAACAGTACATTCAGTTGTTGCTGACGACAGATTAATCTGTGGAACCTTGA 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTCAGGCAGAAAATCCTTTTGTATTTTATGGAACACATTTCTTTAGAAGGAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAAAATTTCTTTCAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGCAGAAACC 1020
Db 321 ThrAsnPheGluGlyLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGTTTCCACCTTGGATCGAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
```

```
Db 21 AspThrAsnGluSerGluLeuLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
QY 121 CGTTTGTGTCATCTTCCCAATCCAGTACCCTGATATTTGGAAATGTATAAACAAGGCACAG 180
Db 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
QY 181 GCTTCTCTGACAGCAGAGAGAGTGCAGCTTATCAAGAGTCTCCCTCAGTGGAAACAAG 240
Db 61 AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrPheAsnLys 80
QY 241 CTTAAGCAGATGAGAAGTACTCTCATCTCATCTAGCTTTTTCAGCCAGTGTAT 300
Db 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
QY 301 CGAATTTGTAATGAAATTTGGTGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
Db 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
QY 361 CGCTGTTTCTATGGCTTTCAAAATTCATCGAGAATGTTTCACTCAGAGATGTACAGTTTG 420
Db 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetPheSerLeu 140
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTTGAA 480
Db 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
QY 481 ACCATGCCCTATGTTTAAAGAAAACAGATTCGGCTTCGAGTGTATGATGATGACATAGAAA 540
Db 161 ThrMetProTyrValLysLysLysAlaAspTyrAlaLeuArgTyrPheAlaAspArgLys 180
QY 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTGTAGAGAGATTTTCTTCTCAGGA 600
Db 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
QY 601 TCTTTTGTCTATATCTGGCTTAAAGAGAGAGTCTTATCCAGGACTCACTTTTTC 660
Db 201 SerPheAlaAlaIlePheTyrLysLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
QY 661 AATGACTCATCAGCAGATGAGGACTTCACTGTGCTAGAGTCTTGTGCTGATGTTCCAA 720
Db 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
QY 721 TACTTAGTAAATAGCTTCAAGAAAAGGCTTCAGAGAGTCTTATGGAATGCAATTTGCATT 840
Db 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleValAspAlaValLys 260
QY 781 ATGAGCAGGAGTTTAAACAGAACCTTGCAGTTCGCTCATTCAGTGAATGCAATTTGCATT 840
Db 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAACAGTACATTGAGTTGTAGCTGACAGATTAATGTTGGAACCTTGATTTCTCA 900
Db 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTCAGGAGAGAAAATCTTTTCAATTTATGGAACATTTCTTTAGAGGAAAA 960
Db 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluLys 320
QY 961 ACAAAATTTCTTTGAAACAGAGTTTCAGAGTATCAGCGTTTTCAGTGTATGCGAGAAACC 1020
Db 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGCTTCACCTTGATGTCAGATTTT 1053
Db 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
```

RESULT 15

PCT-US05-09639-27

```
; Sequence 27, Application pc/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004WO1
```

```
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
; PCT-US05-09639-27
```

Alignment Scores:

Pred. No.:	5,43e-177	Length:	351
Score:	1811.00	Matches:	349
Percent Similarity:	100.00%	Conservative:	2
Best Local Similarity:	99.43%	Mismatches:	0
Query Match:	98.10%	Indels:	0
DB:	1	Gaps:	0

US-10-698-228-2 (1-1053) x PCT-US05-09639-27 (1-351)

QY	1	ATGGCGACCCGAAAGCGCGGCGGCTGGATCAGATGAGAGATCATCTTTCA	60
Db	1	MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer	20
QY	61	GACACCAACGAAGTGAATTAAGTCAATGAAGAGCCACTCCTTAAGAAAGAGTCTCGC	120
Db	21	AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSer	40
QY	121	CGGTTTGTCTATCTTCCAAATCCAGTACCTCGATATTTGGAAAAATGTATAAACAGGCACAG	180
Db	41	ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln	60
QY	181	GCTTCTCTTGGACAGCAGAGGTCGATTAATCAAAAGGATCTCCCTCAGTGGAAACAAG	240
Db	61	AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrPheAsnLys	80
QY	241	CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCATCTTAGCCTTTTTCAGCCAGCGATGAT	300
Db	81	LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp	100
QY	301	GGAATTTGTAATGAAATTTGGTGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT	360
Db	101	GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla	120
QY	361	CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAGAATTTTCACTCAGAGATGTACAGTTTG	420
Db	121	ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrPheSerLeu	140
QY	421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTTGAA	480
Db	141	LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu	160
QY	481	ACCATGCCCTATGTTTAAAGAAAACAGATTCGGCTTCGAGTGTATGATGATGACATAGAAA	540
Db	161	ThrMetProTyrValLysLysLysAlaAspTyrAlaLeuArgTyrPheAlaAspArgLys	180
QY	541	TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTGTAGAGAGATTTTCTTCTCAGGA	600
Db	181	SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly	200
QY	601	TCTTTTGTCTATATTTGGCTTAAAGAGAGAGTCTTATCCAGGACTCACTTTTTC	660
Db	201	SerPheAlaAlaIlePheTyrLysLysLysArgGlyLeuMetProGlyLeuThrPheSer	220
QY	661	AATGACTCATCAGCAGATGAGGACTTCACTGTGCTAGAGTCTTGTGCTGATGTTCCAA	720
Db	221	AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln	240
QY	721	TACTTAGTAAATAGCTTCAAGAAAAGGCTTCAGAGAGTCTTATGGAATGCAATTTGCATT	780
Db	241	TyrLeuValAsnLysProSerGluGluArgValArgGluIleValAspAlaValLys	260
QY	781	ATGAGCAGGAGTTTAAACAGAACCTTGCAGTTCGCTCATTCAGTGAATGCAATTTGCATT	840
Db	261	IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle	280
QY	841	TTGATGAACAGTACATTGAGTTGTAGCTGACAGATTAATGTTGGAACCTTGATTTCTCA	900
Db	281	LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer	300
QY	901	AAGGTTTTTCAGGAGAGAAAATCTTTTCAATTTATGGAACATTTCTTTAGAGGAAAA	960
Db	301	LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluLys	320
QY	961	ACAAATTTCTTTGAAACAGAGTTTCAGAGTATCAGCGTTTTCAGTGTATGCGAGAAACC	1020
Db	321	ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr	340
QY	1021	ACAGATAACGCTTCACCTTGATGTCAGATTTT	1053
Db	341	ThrAspAsnValPheThrLeuAspAlaAspPhe	351

QY 781 ATTGAGCAGGAGTTTAAACAGAACCTTGCAGTTGGCTCATTGGAATGAATTGCATT 840
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
QY 841 TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGAACTTGGATTCTCA 900
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
QY 901 AAGGTTTTTCAGGCAGAAATCCTTTTGATTTTATGGAACATTCTTTAGAAGGAAA 960
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
QY 961 ACAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGCAGAAACC 1020
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
321 ThrAsnPhePheGluLysArgValSerGluTyGlnArgPheAlaValMetAlaGluThr 340
QY 1021 ACAGATAACGTTCTTCCACCTTGGATGCAGATTTT 1053
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
341 ThrAsnValPheThrLeuAspAlaAspPhe 351

Search completed: October 30, 2005, 06:50:27
Job time : 43 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 30, 2005, 00:02:51 ; Search time 1372.5 Seconds
(without alignments)
11030.942 Million cell updates/sec

Title: US-10-698-228-2
Perfect score: 1053
Sequence: 1 atggggcagccgaaagcc.....tcacctgtgacagatttt 1053

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 27723106 seqs, 7188968421 residues
Total number of hits satisfying chosen parameters: 55446212

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents NA New:
1: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:
2: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:
3: /cgn2_6/ptodata/1/pna/US05_NEW_COMB.seq:
4: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq:
5: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq:
6: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:
7: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:
8: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:
9: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:
10: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:
11: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:
12: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:
13: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:
14: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:
15: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:
16: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:
17: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:
18: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:
19: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:
20: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq:
21: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq:
22: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq:
23: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq:
24: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq:
25: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq:
26: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq:
27: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq:
28: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					Description	
Result No.	Score	Query Match	Length	ID		
1	1053	100.0	1056	1	PCT-US05-09639-16	Sequence 16, Appl
2	1053	100.0	1056	27	US-11-088-686-16	Sequence 16, Appl
3	1053	100.0	4929	1	PCT-US05-22501-1545	Sequence 1545, Ap
4	1053	100.0	4955	1	PCT-US05-09639-44	Sequence 44, Appl
5	1053	100.0	4955	1	PCT-US05-21650-13	Sequence 13, Appl

6	1053	100.0	4955	27	US-11-088-686-44	Sequence 44, Appl
7	1052.6	100.0	1849	15	US-10-990-328A-49	Sequence 49, Appl
8	1051.4	99.8	1056	1	PCT-US05-09639-18	Sequence 18, Appl
9	1051.4	99.8	1056	27	US-11-088-686-18	Sequence 18, Appl
10	1049.8	99.7	1056	1	PCT-US05-09639-20	Sequence 20, Appl
11	1049.8	99.7	1056	27	US-11-088-686-20	Sequence 20, Appl
12	1048.2	99.5	1056	1	PCT-US05-09639-22	Sequence 22, Appl
13	1048.2	99.5	1056	1	PCT-US05-09639-24	Sequence 24, Appl
14	1048.2	99.5	1056	27	US-11-088-686-22	Sequence 22, Appl
15	1048.2	99.5	1056	27	US-11-088-686-24	Sequence 24, Appl
16	1046.6	99.4	1056	1	PCT-US05-09639-26	Sequence 26, Appl
17	1046.6	99.4	1056	1	PCT-US05-09639-28	Sequence 28, Appl
18	1046.6	99.4	1056	27	US-11-088-686-26	Sequence 26, Appl
19	1046.6	99.4	1056	27	US-11-088-686-28	Sequence 28, Appl
20	624	59.3	624	1	PCT-US05-09639-42	Sequence 42, Appl
21	624	59.3	624	27	US-11-088-686-42	Sequence 42, Appl
22	578.4	54.9	1170	1	PCT-US05-09639-2	Sequence 2, Appl
23	578.4	54.9	1170	27	US-11-088-686-2	Sequence 2, Appl
24	578.4	54.9	1649	9	US-10-553-098-583	Sequence 583, App
25	578.4	54.9	1649	13	US-10-955-054A-135	Sequence 135, App
26	578.4	54.9	2470	1	PCT-US05-13260-66	Sequence 66, Appl
27	578.4	54.9	2470	1	PCT-US05-20840-66	Sequence 66, Appl
28	578.4	54.9	2470	1	PCT-US05-22501-6487	Sequence 6487, Ap
29	578.4	54.9	2479	15	US-10-940-774A-2025	Sequence 2025, Ap
30	578.4	54.9	2482	11	US-10-302-689A-107255	Sequence 107255,
31	578.4	54.9	2500	1	PCT-US05-09639-43	Sequence 43, Appl
32	578.4	54.9	2500	1	PCT-US05-20435-12	Sequence 12, Appl
33	578.4	54.9	2500	1	PCT-US05-21650-12	Sequence 12, Appl
34	578.4	54.9	2500	1	PCT-US05-27243-78	Sequence 78, Appl
35	578.4	54.9	2500	14	US-10-960-414-154	Sequence 154, App
36	578.4	54.9	2500	15	US-10-940-774A-145	Sequence 145, App
37	578.4	54.9	2500	24	US-11-000-688A-1211	Sequence 1211, Ap
38	578.4	54.9	2500	24	US-11-150-888-12	Sequence 12, Appl
39	578.4	54.9	2500	27	US-11-088-686-43	Sequence 43, Appl
40	578	54.9	2909	15	US-10-990-328A-4852	Sequence 4852, Ap
41	578	54.9	3616	15	US-10-990-328A-4851	Sequence 4851, Ap
42	576.8	54.8	1170	1	PCT-US05-09639-4	Sequence 4, Appli
43	576.8	54.8	1170	27	US-11-088-686-4	Sequence 4, Appli
44	575.2	54.6	1170	1	PCT-US05-09639-6	Sequence 6, Appli
45	575.2	54.6	1170	1	PCT-US05-09639-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
PCT-US05-09639-16
; Sequence 16, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1056)
PCT-US05-09639-16

Query Match 100.0%; Score 1053; DB 1; Length 1056;
Best Local Similarity 100.0%; Pred. No. 7,1e-271;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGCGACCCGAAAGCGGAGCGCGCGCTGATCAGATGAGATCATCTTCA 60
|||||

Db 1 ATGGGCGACCCGGAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
QY 61 GACACCAACGAAGTGAATAAGTCAAAATGAAGACCACTCTTAAGAAGAGTTCTCGC 120
Db 61 GACACCAACGAAGTGAATAAGTCAAAATGAAGACCACTCTTAAGAAGAGTTCTCGC 120
QY 121 CGGTTTGTGCATCTTTCCAAATCCAGTACCTCGATATTTGGAAAATGTATAAACAGGCACAG 180
Db 121 CGGTTTGTGCATCTTTCCAAATCCAGTACCTCGATATTTGGAAAATGTATAAACAGGCACAG 180
QY 181 GCTTCCTCTGACAGCAGAGAGGTCGATCTATCAAGAGATCTCCCTCACTGGAAACAAG 240
Db 181 GCTTCCTCTGACAGCAGAGAGGTCGATCTATCAAGAGATCTCCCTCACTGGAAACAAG 240
QY 241 CTTAAAGCAGATGAGAAGTACCTCTCATCTCATCTTACCTTGGAGGCTTTTTCGACCCAGTGAT 300
Db 241 CTTAAAGCAGATGAGAAGTACCTCTCATCTCATCTTACCTTGGAGGCTTTTTCGACCCAGTGAT 300
QY 301 GGAAATTTGTAATGAAAATTTGGTGGAGCGCTTTTAGTCAGGAGGTGAGGTTCCAGAGGCT 360
Db 301 GGAAATTTGTAATGAAAATTTGGTGGAGCGCTTTTAGTCAGGAGGTGAGGTTCCAGAGGCT 360
QY 361 CGCTGTTTCTATGGCTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGGCTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
QY 421 CTGATAGACACTTACATCAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATGAA 480
Db 421 CTGATAGACACTTACATCAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATGAA 480
QY 481 ACCATGCCCTATGTTAAGAAAAGGGAATTTTGGTGGAGCGCTTTTAGTCAGGAGGTGAGGTTCCAGAGGCT 540
Db 481 ACCATGCCCTATGTTAAGAAAAGGGAATTTTGGTGGAGCGCTTTTAGTCAGGAGGTGAGGTTCCAGAGGCT 540
QY 541 TCTACTTTTGGGAAAAGAGTGTGGCTTTGCTGTGTAGAGAGGAGTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAAGAGTGTGGCTTTGCTGTGTAGAGAGGAGTCTTCTCAGGA 600
QY 601 TCTTTTGTGCTATATTTCTGGCTTAAGAGAGAGGTCCTTATGCCAGGACTCACTTTTCC 660
Db 601 TCTTTTGTGCTATATTTCTGGCTTAAGAGAGAGGTCCTTATGCCAGGACTCACTTTTCC 660
QY 661 AATGAACACTCATCAGCAGATGAAGAGCTTCACTGTGACCTTTGCTGCTGATGTTCCAA 720
Db 661 AATGAACACTCATCAGCAGATGAAGAGCTTCACTGTGACCTTTGCTGCTGATGTTCCAA 720
QY 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGCTCAGGAGATCATTTGTTGATGCTGCTCAA 780
Db 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGCTCAGGAGATCATTTGTTGATGCTGCTCAA 780
QY 781 ATTGAGCAGAGTTTAAACAGAGCCTTGGCCAGTTGGGCTCATTTGGAATGAATTGCAAT 840
Db 781 ATTGAGCAGAGTTTAAACAGAGCCTTGGCCAGTTGGGCTCATTTGGAATGAATTGCAAT 840
QY 841 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTGTGGAATCTTGATTTCTCA 900
Db 841 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTGTGGAATCTTGATTTCTCA 900
QY 901 AAGGTTTTTTCAGGAGAAAATCCTTTTGTATTTTATGGAACAACTTTCTTTAGAGGAAAA 960
Db 901 AAGGTTTTTTCAGGAGAAAATCCTTTTGTATTTTATGGAACAACTTTCTTTAGAGGAAAA 960
QY 961 ACAATATTTCTTTGAGAAAAGGTTTTCAGAGTATCAGCGTTTTCAGGATTTGCGAGAAACC 1020
Db 961 ACAATATTTCTTTGAGAAAAGGTTTTCAGAGTATCAGCGTTTTCAGGATTTGCGAGAAACC 1020
QY 1021 ACAGATAACGCTTTCACCTTGGATGCGAGATTTT 1053
Db 1021 ACAGATAACGCTTTCACCTTGGATGCGAGATTTT 1053

; Sequence 16, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; PRIOR FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1056)
US-11-088-686-16

Query Match 100.0%; Score 1053; DB 27; Length 1056;
Best Local Similarity 100.0%; Pred. No. 7.1e-271;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAGCGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 ATGGGCGACCCGGAAGCGCGGAGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
QY 61 GACACCAACGAAGTGAATAAGTCAAAATGAAGAGCCTCTCTAAGAAAAGTTCGCG 120
Db 61 GACACCAACGAAGTGAATAAGTCAAAATGAAGAGCCTCTCTAAGAAAAGTTCGCG 120
QY 121 CGGTTTGTGCATCTTTCCAAATCCAGTACCTCGATATTTGGAAAATGTATAAACAGGCACAG 180
Db 121 CGGTTTGTGCATCTTTCCAAATCCAGTACCTCGATATTTGGAAAATGTATAAACAGGCACAG 180
QY 181 GCTTCCTCTGACAGCAGAGAGGTCGATCTATCAAGAGATCTCCCTCACTGGAAACAAG 240
Db 181 GCTTCCTCTGACAGCAGAGAGGTCGATCTATCAAGAGATCTCCCTCACTGGAAACAAG 240
QY 241 CTTAAAGCAGATGAGAAGTACCTCTCATCTCTCATCTTACCTTGGAGGCTTTTTCGAGCCAGTGAT 300
Db 241 CTTAAAGCAGATGAGAAGTACCTCTCATCTCTCATCTTACCTTGGAGGCTTTTTCGAGCCAGTGAT 300
QY 301 GGAAATTTGTAATGAAAATTTGGTGGAGCGCTTTTAGTCAGGAGGTGAGGTTCCAGAGGCT 360
Db 301 GGAAATTTGTAATGAAAATTTGGTGGAGCGCTTTTAGTCAGGAGGTGAGGTTCCAGAGGCT 360
QY 361 CGCTGTTTCTATGGCTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGGCTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAATGCAATTTGAA 480
QY 481 ACCATGCCCTATGTTAAGAAAAGGGAATTTTGGTGGAGCGCTTTTAGTCAGGAGGTGAGGTTCCAGAGGCT 540
Db 481 ACCATGCCCTATGTTAAGAAAAGGGAATTTTGGTGGAGCGCTTTTAGTCAGGAGGTGAGGTTCCAGAGGCT 540
QY 541 TCTACTTTTGGGAAAAGAGTGTGGCTTTGCTGTGTAGAGAGGAGTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAAGAGTGTGGCTTTGCTGTGTAGAGAGGAGTCTTCTCAGGA 600
QY 601 TCTTTTGTGCTATATTTCTGGCTTAAGAGAGAGGTCCTTATGCCAGGACTCACTTTTCC 660
Db 601 TCTTTTGTGCTATATTTCTGGCTTAAGAGAGAGGTCCTTATGCCAGGACTCACTTTTCC 660
QY 661 AATGAACACTCATCAGCAGAGATGAAGAGCTTCACTGTGACCTTTGCTGCTGATGTTCCAA 720
Db 661 AATGAACACTCATCAGCAGAGATGAAGAGCTTCACTGTGACCTTTGCTGCTGATGTTCCAA 720
QY 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGCTCAGGAGATCATTTGTTGATGCTGCTCAA 780


```
Db 721 TACCTAGTAATAGCCCTTCAAGAAAGGGTTCAGGAGATCATTTGTTGATGCTGTCAA 780
Qy 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTTGGCTCATTTGGAATGAATTCGATT 840
Db 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGCAGTTGGCTCATTTGGAATGAATTCGATT 840
Qy 841 TTGATGAACAGTACATAGGTTTGTAGCTGACAGATTAATTTGTGGAACCTTGGATTCTCA 900
Db 841 TTGATGAACAGTACATAGGTTTGTAGCTGACAGATTAATTTGTGGAACCTTGGATTCTCA 900
Qy 901 AAGGTTTTTCAGGCGAGAAATCCTTTGATTTATGGAATAAATTTTGAAGAGAAA 960
Db 901 AAGGTTTTTCAGGCGAGAAATCCTTTGATTTATGGAATAAATTTTGAAGAGAAA 960
Qy 961 ACAAAATTTCTTGGAGAACAGTTTCAGAGTATCAGCGTTTGGCAGTTATGCGAGAAACC 1020
Db 961 ACAAAATTTCTTGGAGAACAGTTTCAGAGTATCAGCGTTTGGCAGTTATGCGAGAAACC 1020
Qy 1021 ACAGATAACGCTCTTACCTTGGATGAGATTTT 1053
Db 1021 ACAGATAACGCTCTTACCTTGGATGAGATTTT 1053

RESULT 3
PCT-US05-22501-1545
; Sequence 1545, Application PC/TUS0522501
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals
; TITLE OF INVENTION: Determining Cancer-Linked Gene and Therapeutic Targets Using
; FILE OF INVENTION: Molecular Cytogenetic Methods
; CURRENT APPLICATION NUMBER: PCT/US05/22501
; PRIOR FILING DATE: 2005-07-07
; PRIOR FILING DATE: 2004-06-23
; NUMBER OF SEQ ID NOS: 7840
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1545
; LENGTH: 4929
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: cDNA sequence
PCT-US05-22501-1545

Query Match 100.0%; Score 1053; DB 1; Length 4929;
Best Local Similarity 100.0%; Pred. No. 1.1e-270;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCGGACCCGGAAGCGCGGAGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 245 ATGGCGGACCCGGAAGCGCGGAGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
Qy 61 GACACCAACGGAAGTGAATAAGTCAATGAAGGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 305 GACACCAACGGAAGTGAATAAGTCAATGAAGGCCACTCTTAAGAAAGAGTTCTCGC 364
Qy 121 CGTTTGTTCATCTTCCATCCAGTACCTGATATTTGGAAAAATGATAAAGAGGACAG 180
Db 365 CGTTTGTTCATCTTCCATCCAGTACCTGATATTTGGAAAAATGATAAAGAGGACAG 424
Qy 181 GCTTCCTTCTGACAGCAGAGAGTGCAGTTATCAAAAGGATCTCCCTCACTGGAAACAG 240
Db 425 GCTTCCTTCTGACAGCAGAGAGTGCAGTTATCAAAAGGATCTCCCTCACTGGAAACAG 484
Qy 241 CTTAAGCAGATGAGAAGTCTTCACTCTCATCTAGCTTTTTCAGCCAGTGTAT 300
Db 485 CTTAAGCAGATGAGAAGTCTTCACTCTCATCTAGCTTTTTCAGCCAGTGTAT 544
Qy 301 GGAATTTGAAATGAAAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCGAGGTCCAGAGCT 360
Db 545 GGAATTTGAAATGAAAAATTTGGTGAGCGCTTTAGTCAGGAGGTGCGAGGTCCAGAGCT 604
```

```
Qy 361 CGCTGTTTCTATGGCTTTCAAAATTTCTATCGAAGATTTTCACTCAGAGATGTACAGTTTG 420
Db 605 CGCTGTTTCTATGGCTTTCAAAATTTCTATCGAAGATTTTCACTCAGAGATGTACAGTTTG 664
Qy 421 CTGATAGACACTTATCATCAGAGATCCCAAGAAAGGAAATTTTATTAATGAATGAAGAA 480
Db 665 CTGATAGACACTTATCATCAGAGATCCCAAGAAAGGAAATTTTATTAATGAATGAAGAA 724
Qy 481 ACCATGCCCTATGTTAAGAAAGAGAGATTTGGGCTTGGATGGATGAGATAGAGAAA 540
Db 725 ACCATGCCCTATGTTAAGAAAGAGAGATTTGGGCTTGGATGGATGAGATAGAGAAA 784
Qy 541 TCTACTTTTGGGAGAAAGAGTGGTGGCTTTGCTGTGTAAGAGAGTGTTCCTCTCAGGA 600
Db 785 TCTACTTTTGGGAGAAAGAGTGGTGGCTTTGCTGTGTAAGAGAGTGTTCCTCTCAGGA 844
Qy 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGTCTTATGCCAGAGTCACTTTTTC 660
Db 845 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGTCTTATGCCAGAGTCACTTTTTC 904
Qy 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 905 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 964
Qy 721 TACTTTAGTAATAAGCCTTTCAGAAAGAGAGGTTCAGGAGATCATTTGTTGATGCTGTCAA 780
Db 965 TACTTTAGTAATAAGCCTTTCAGAAAGAGAGGTTCAGGAGATCATTTGTTGATGCTGTCAA 1024
Qy 781 ATTGAGCAGAGTGTTTTAAACAGAAAGCCTTGCAGTGTGGCTCATTTGGAATGAATTCATT 840
Db 1025 ATTGAGCAGAGTGTTTTAAACAGAAAGCCTTGCAGTGTGGCTCATTTGGAATGAATTCATT 1084
Qy 841 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTTGTGAACTTTGGAATTCCTCA 900
Db 1085 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTAATCTTGTGAACTTTGGAATTCCTCA 1144
Qy 901 AAGGTTTTTTCAGCAGAGAAATCCTTTTGTATTTTATGGAACACATTTCTTTAGAGGAAA 960
Db 1145 AAGGTTTTTTCAGCAGAGAAATCCTTTTGTATTTTATGGAACACATTTCTTTAGAGGAAA 1204
Qy 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGATATCAGCGTTTTCAGTATGCGAGAAACC 1020
Db 1205 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGATATCAGCGTTTTCAGTATGCGAGAAACC 1264
Qy 1021 ACAGATAACGCTCTTCACTTGGATGAGATTTT 1053
Db 1265 ACAGATAACGCTCTTCACTTGGATGAGATTTT 1297

RESULT 4
PCT-US05-09639-44
; Sequence 44, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; PRIOR FILING DATE: 2005-03-23
; PRIOR FILING DATE: 2005-03-23
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US05-09639-44

Query Match 100.0%; Score 1053; DB 1; Length 4955;
Best Local Similarity 100.0%; Pred. No. 1.1e-270;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 ATGGCGACCCCGAAAGCGCGGAGCGGCTGGATCAGAGTGAAGATCATCTTCA 60
Db 245 ATGGCGACCCCGAAAGCGCGGAGCGGCTGGATCAGAGTGAAGATCATCTTCA 304
QY 61 GACACCAAGAGTGAATTAAGTCAATTAAGAGCCACTCTTAAGAAAGTCTTCGC 120
Db 305 GACACCAAGAGTGAATTAAGTCAATTAAGAGCCACTCTTAAGAAAGTCTTCGC 364
QY 121 CGGTTGTCTATCTTCCAAATCAGTACCTGATATTTGGAAAATGTATAAAGAGGACAG 180
Db 365 CGGTTGTCTATCTTCCAAATCAGTACCTGATATTTGGAAAATGTATAAAGAGGACAG 424
QY 181 GCTTCCTCTGACAGCAGAGAGTGCATTTCAATCAAGAGTCTCCCTCACTGGAAACAG 240
Db 425 GCTTCCTCTGACAGCAGAGAGTGCATTTCAATCAAGAGTCTCCCTCACTGGAAACAG 484
QY 241 CTTAAAGCAGATGAGAGTACTTCTCATCTCATCTTAAAGAGTCTTCTGAGCAGTGTAT 300
Db 485 CTTAAAGCAGATGAGAGTACTTCTCATCTCATCTTAAAGAGTCTTCTGAGCAGTGTAT 544
QY 301 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTCCAGAGGCT 360
Db 545 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTCCAGAGGCT 604
QY 361 CGCTGTTCTATGCTTTCAAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 420
Db 605 CGCTGTTCTATGCTTTCAAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 664
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTAAATGCAATTTGAA 480
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTAAATGCAATTTGAA 724
QY 481 ACCATGCCCTATGTTAAGAAAAGAGAGTTCGAGTTCAGAGAGTTCCTCTCACTGAGATGTACAGTTTG 540
Db 725 ACCATGCCCTATGTTAAGAAAAGAGAGTTCGAGTTCAGAGAGTTCCTCTCACTGAGATGTACAGTTTG 784
QY 541 TCTACTTTTGGGAAAAGAGTGGTGGGCTTTGCTGCTGTAGAAAAGGAAATTTTAAATGCAATTTGAA 600
Db 785 TCTACTTTTGGGAAAAGAGTGGTGGGCTTTGCTGCTGTAGAAAAGGAAATTTTAAATGCAATTTGAA 844
QY 601 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCGAGACACTCATCTTTTCC 660
Db 845 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCGAGACACTCATCTTTTCC 904
QY 661 AATGAACACTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGAGTTCCTCA 720
Db 905 AATGAACACTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGAGTTCCTCA 964
QY 721 TACTTAGTAAATAAGCCTTCAGAAAAGAGGTTCAGGAGATCATTTGTTGATGCTGTCAAA 780

RESULT 5
PCT-US05-21650-13
; Sequence 13, Application PC/TUS0521650
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: MIGFS AS MODIFIERS OF THE IGF PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX05-017
; CURRENT APPLICATION NUMBER: PCT/US05/21650
; CURRENT FILING DATE: 2005-06-27
; PRIOR APPLICATION NUMBER: 60/581,689
; PRIOR FILING DATE: 2004-06-21
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US05-21650-13

Query Match 100.0%; Score 1053; DB 1; Length 4955;
Best Local Similarity 100.0%; Pred. No. 1.1e-270;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGACCCCGAAAGCGGAGCGGCTGGATCAGAGTGAAGATCATCTTCA 60
Db 245 ATGGCGACCCCGAAAGCGGAGCGGCTGGATCAGAGTGAAGATCATCTTCA 304
QY 61 GACACCAAGAGTGAATTAAGTCAATTAAGAGCCACTCTTAAGAAAGTCTTCGC 120
Db 305 GACACCAAGAGTGAATTAAGTCAATTAAGAGCCACTCTTAAGAAAGTCTTCGC 364
QY 121 CGGTTGTCTATCTTCCAAATCAGTACCTGATATTTGGAAAATGTATAAAGAGGACAG 180
Db 365 CGGTTGTCTATCTTCCAAATCAGTACCTGATATTTGGAAAATGTATAAAGAGGACAG 424
QY 181 GCTTCCTCTGACAGCAGAGAGTGCATTTCAATCAAGAGTCTCCCTCACTGAGAAACAG 240
Db 425 GCTTCCTCTGACAGCAGAGAGTGCATTTCAATCAAGAGTCTCCCTCACTGAGAAACAG 484
QY 241 CTTAAAGCAGATGAGAGTACTTCTCATCTCATCTTAAAGAGTCTTCTGAGCAGTGTAT 300
Db 485 CTTAAAGCAGATGAGAGTACTTCTCATCTCATCTTAAAGAGTCTTCTGAGCAGTGTAT 544
QY 301 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTCCAGAGGCT 360
Db 545 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTCCAGAGGCT 604
QY 361 CGCTGTTCTATGCTTTCAAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 420
Db 605 CGCTGTTCTATGCTTTCAAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 664
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTAAATGCAATTTGAA 480
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTAAATGCAATTTGAA 724
QY 481 ACCATGCCCTATGTTAAGAAAAGAGAGTTCGAGTTCAGAGAGTTCCTCTCACTGAGATGTACAGTTTG 540
Db 725 ACCATGCCCTATGTTAAGAAAAGAGAGTTCGAGTTCAGAGAGTTCCTCTCACTGAGATGTACAGTTTG 784
QY 541 TCTACTTTTGGGAAAAGAGTGGTGGGCTTTGCTGCTGTAGAAAAGGAAATTTTAAATGCAATTTGAA 600
Db 785 TCTACTTTTGGGAAAAGAGTGGTGGGCTTTGCTGCTGTAGAAAAGGAAATTTTAAATGCAATTTGAA 844
QY 601 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCGAGACACTCATCTTTTCC 660
Db 845 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCGAGACACTCATCTTTTCC 904
QY 661 AATGAACACTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGAGTTCCTCA 720
Db 905 AATGAACACTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGCCTGAGTTCCTCA 964
QY 721 TACTTAGTAAATAAGCCTTCAGAAAAGAGGTTCAGGAGATCATTTGTTGATGCTGTCAAA 780

Db 965 TACTTAGTAAATAAGCCTTCAGAGAAAGGGTCAGGAGATCATTTGTGCTGTCAAA 1024
QY 781 ATTGACGAGGTTTTTAACAGAGCCTTGGCAGTTGGCTCATTTGGAATGCAATT 840
Db 1025 ATTGACGAGGTTTTTAACAGAGCCTTGGCAGTTGGCTCATTTGGAATGCAATT 1084
QY 841 TTGATGAACAGTACATTTGAGTTTGTAGTGACAGATTACTTTGTGGAACCTTGGATTCTCA 900
Db 1085 TTGATGAACAGTACATTTGAGTTTGTAGTGACAGATTACTTTGTGGAACCTTGGATTCTCA 1144
QY 901 AAGGTTTTTTCAGGAGAAATCCTTTTGAATTTATGGAATAAATTTCTTTGAAGAAAA 960
Db 1145 AAGGTTTTTTCAGGAGAAATCCTTTTGAATTTATGGAATAAATTTCTTTGAAGAAAA 1204
QY 961 ACAAAATTTCTTCAGAAACGAGTTTCAGAGTATCAGCGTTTGTGAGTTATGCGAGAAAC 1020
Db 1205 ACAAAATTTCTTCAGAAACGAGTTTCAGAGTATCAGCGTTTGTGAGTTATGCGAGAAAC 1264
QY 1021 ACAGATAACGTTCTTCACTTGGATGAGATTTT 1053
Db 1265 ACAGATAACGTTCTTCACTTGGATGAGATTTT 1297

RESULT 6
US-11-088-686-44
; Sequence 44, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-088-686-44

Query Match 100.0%; Score 1053; DB 27; Length 4955;
Best Local Similarity 100.0%; Pred. No. 1.1e-270; Indels 0; Gaps 0;
Matches 1053; Conservative 0; Mismatches 0;

QY 1 ATGGGCGACCCGGAAGCGGAGCGGCGGCTGGATCAGATCAGAGATCATCTTCA 60
Db 245 ATGGGCGACCCGGAAGCGGAGCGGCGGCTGGATCAGATCAGAGATCATCTTCA 304
QY 61 GACACCAACGAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 305 GACACCAACGAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 364
QY 121 CGTTTGTTCATCTTCCATCAGTACCTGTATTTGGAATGTATTAACAGGACAG 180
Db 365 CGTTTGTTCATCTTCCATCAGTACCTGTATTTGGAATGTATTAACAGGACAG 424
QY 181 GCTTCTCTTCGACAGCAGAGAGGTCGACTTATCAAGGATCTCCCTCACTGGAACAG 240
Db 425 GCTTCTCTTCGACAGCAGAGAGGTCGACTTATCAAGGATCTCCCTCACTGGAACAG 484
QY 241 CTTAAAGCAGATGAGAAGTACTTCTCTCACTTTAGCCCTTTTTCAGCCAGTGTAT 300
Db 485 CTTAAAGCAGATGAGAAGTACTTCTCTCACTTTAGCCCTTTTTCAGCCAGTGTAT 544
QY 301 GGAATTTGAATGAATAATTTGTGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGT 360
Db 545 GGAATTTGAATGAATAATTTGTGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGT 604
QY 361 CGCTGTTCTATGGCTTTTCAAAATCTTCATCAGAGATGTTTCACTCAGAGATGTACAGTTTG 420
Db 605 CGCTGTTCTATGGCTTTTCAAAATCTTCATCAGAGATGTTTCACTCAGAGATGTACAGTTTG 664

QY 421 CTGATAGACACTTTACATCAGAGATCCCAAGAAAAAGGAAATTTTATTAATGCAATTGAA 480
Db 665 CTGATAGACACTTTACATCAGAGATCCCAAGAAAAAGGAAATTTTATTAATGCAATTGAA 724
QY 481 ACCATGCCCTATCTTAAGAAAAAGCAGATTGGGCCCTTGGCATGATAGAGATAGAAAA 540
Db 725 ACCATGCCCTATCTTAAGAAAAAGCAGATTGGGCCCTTGGCATGATAGAGATAGAAAA 784
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCCCTTGTCTGCTGTAGAGAGTTTCTTCTCAGGA 600
Db 785 TCTACTTTTGGGGAAGAGTGGTGGCCCTTGTCTGCTGTAGAGAGTTTCTTCTCAGGA 844
QY 601 TCTTTTGTCTATATTTCTGGCTAAAGAGAGGCTTATGCCAGGACTCACATTTTTC 660
Db 845 TCTTTTGTCTATATTTCTGGCTAAAGAGAGGCTTATGCCAGGACTCACATTTTTC 904
QY 661 AATGAACCTCATCAGCAGAGATGAAGACTTTCACCTGTGACTTTGTCTGCTGATGTTCCAA 720
Db 905 AATGAACCTCATCAGCAGAGATGAAGACTTTCACCTGTGACTTTGTCTGCTGATGTTCCAA 964
QY 721 TACTTAGTAAATAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAAA 780
Db 965 TACTTAGTAAATAGCCTTCAGAGAAAGGCTCAGGAGATCATTTGTGATGCTGTCAAA 1024
QY 781 ATTGACGAGGTTTTTAAACAGAGCCTTGGCAGTTGGCTCATTTGGAATGCAATTGCAAT 840
Db 1025 ATTGACGAGGTTTTTAAACAGAGCCTTGGCAGTTGGCTCATTTGGAATGCAATTGCAAT 1084
QY 841 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTACTTTGTGGAACCTTGGATTCTCA 900
Db 1085 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTACTTTGTGGAACCTTGGATTCTCA 1144
QY 901 AAGGTTTTTTCAGGAGAAATCCTTTTGAATTTTATGGAATAAATTTCTTTTGAAGAAAA 960
Db 1145 AAGGTTTTTTCAGGAGAAATCCTTTTGAATTTTATGGAATAAATTTCTTTTGAAGAAAA 1204
QY 961 ACAAAATTTCTTCAGAAACGAGTTTCAGAGTATCAGCGTTTGTGAGTTATGCGAGAAAC 1020
Db 1205 ACAAAATTTCTTCAGAAACGAGTTTCAGAGTATCAGCGTTTGTGAGTTATGCGAGAAAC 1264
QY 1021 ACAGATAACGTTCTTCACTTGGATGAGATTTT 1053
Db 1265 ACAGATAACGTTCTTCACTTGGATGAGATTTT 1297

RESULT 7
US-10-990-328A-49
; Sequence 49, Application US/10990328A
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328A
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 1849
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-990-328A-49

Query Match 100.0%; Score 1052.6; DB 15; Length 1849;
Best Local Similarity 99.9%; Pred. No. 1.1e-270; Indels 0; Gaps 0;
Matches 1052; Conservative 1; Mismatches 0;
QY 1 ATGGGCGACCCGGAAGCGGAGCGGCGGCTGGATCAGAGATGAGATCATCTTCA 60
Db 85 ATGGGCGACCCGGAAGCGGAGCGGCGGCTGGATCAGAGATGAGATCATCTTCA 144

```
QY 61 GACACAAAGTGAATAAGTCAATGAAGAGCCACTCCTAAGAAAGAGTTCTCGC 120
Db 145 GACACAAAGTGAATAAGTCAATGAAGAGCCACTCCTAAGAAAGAGTTCTCGC 204
QY 121 CGGTTTGTCATCTTCCATCCAGTACCTCTGATATTTGGAAATGTATAAAGAGCAG 180
Db 205 CGGTTTGTCATCTTCCATCCAGTACCTCTGATATTTGGAAATGTATAAAGAGCAG 264
QY 181 GCTTCTCTCTGACAGCAGAGAGGTCGACTTATCAAGAGTCTCCTCACTGGAAACAAG 240
Db 265 GCTTCTCTCTGACAGCAGAGAGGTCGACTTATCAAGAGTCTCCTCACTGGAAACAAG 324
QY 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTTAGCCCTTTTTCAGCCAGTGTAT 300
Db 325 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTTAGCCCTTTTTCAGCCAGTGTAT 384
QY 301 GGAATTTGAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGT 360
Db 385 GGAATTTGAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGT 444
QY 361 CGCTGTTTCTATGGCTTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db 445 CGCTGTTTCTATGGCTTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 504
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 505 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 564
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGATGAGATGACAGATAGAAA 540
Db 565 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGATGAGATGACAGATAGAAA 624
QY 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAAGAGGAGTTTCTTCTCAGGA 600
Db 625 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAAGAGGAGTTTCTTCTCAGGA 684
QY 601 TCTTTTGTGCTATATTTCTGGCTAAAGAAAGAGAGTCTTTATGCCAGGACTCACTTTTCC 660
Db 685 TCTTTTGTGCTATATTTCTGGCTAAAGAAAGAGAGTCTTTATGCCAGGACTCACTTTTCC 744
QY 661 AATGAACCTCAGCAGAGATGAAGACTTCTACTGTGACTTTGCTTGGCTGCTGATGTTCCAA 720
Db 745 AATGAACCTCAGCAGAGATGAAGAGTCTACTGTGACTTTGCTTGGCTGCTGATGTTCCAA 804
QY 721 TACTTAGTAAATAAGCCTTCAGAAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 780
Db 805 TACTTAGTAAATAAGCCTTCAGAAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 864
QY 781 ATTGACGAGGAGTTTAAACAGAGCCTTGGCAGTTGGCCTCATTTGGAATGAATTGCATT 840
Db 865 ATTGACGAGGAGTTTAAACAGAGCCTTGGCAGTTGGCCTCATTTGGAATGAATTGCATT 924
QY 841 TTGATGAACAGTACATTTAGTTTGTAGCTGACAGATTAATTGTGGAACCTTGATTTCTCA 900
Db 925 TTGATGAACAGTACATTTAGTTTGTAGCTGACAGATTAATTGTGGAACCTTGATTTCTCA 984
QY 901 AAGGTTTTTCAGGCAGAAATCCTTTGATTTTATGGAATAATTTCTTTAGAGGAAAA 960
Db 985 AAGGTTTTTCAGGCAGAAATCCTTTGATTTTATGGAATAATTTCTTTAGAGGAAAA 1044
QY 961 ACAAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTTATGCGAGAAACC 1020
Db 1045 ACAAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTTATGCGAGAAACC 1104
QY 1021 ACAGATAACGCTTTCACCTTGGATCGAGATTTT 1053
Db 1105 ACAGATAACGCTTTCACCTTGGATCGAGATTTT 1137
```

RESULT 8
PCT-US05-09639-18
; Sequence 18, Application PC/TUS0509639
; GENERAL INFORMATION:

```
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US05-09639-18

Query Match 99.8%; Score 1051.4; DB 1; Length 1056;
Best Local Similarity 99.9%; Pred. No. 1.9e-270;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 ATGGGCGACCCGGAAGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
QY 61 GACACAAAGTGAATAAGTCAATGAAGAGCCACTCCTAAGAAAGAGTTCTCGC 120
Db 61 GACACAAAGTGAATAAGTCAATGAAGAGCCACTCCTAAGAAAGAGTTCTCGC 120
QY 121 CGGTTTGTCATCTTCCATCCAGTACCTGATATTTGGAAATGTATAAAGAGCAG 180
Db 121 CGGTTTGTCATCTTCCATCCAGTACCTGATATTTGGAAATGTATAAAGAGCAG 180
QY 181 GCTTCTCTCTGACAGCAGAGAGGTCGACTTATCAAGAGTCTCCTCACTGGAAACAAG 240
Db 181 GCTTCTCTCTGACAGCAGAGAGGTCGACTTATCAAGAGTCTCCTCACTGGAAACAAG 240
QY 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTTTAGCCCTTTTTCAGCCAGTGTAT 300
Db 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTTTAGCCCTTTTTCAGCCAGTGTAT 300
QY 301 GGAATTTGAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGT 360
Db 301 GGAATTTGAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGT 360
QY 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAATGTTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTTGGCTTTCAAAATCTCATCGAGAATGTTTCACTCAGAGATGTACAGTTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGATGAGATGACAGATAGAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGATGAGATGACAGATAGAAA 540
QY 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAAGAGGAGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGCTGTAAGAGGAGTTTCTTCTCAGGA 600
QY 601 TCTTTTGTGCTATATTTCTGGCTAAAGAAAGAGAGGCTTATGCGAGGACTCACCTTTTTC 660
Db 601 TCTTTTGTGCTATATTTCTGGCTAAAGAAAGAGAGGCTTATGCGAGGACTCACCTTTTTC 660
QY 661 AATGAACCTCAGCAGAGATGAAGAGTCTCACTGTGACTTTGCTTGGCTGCTGATGTTCAA 720
Db 661 AATGAACCTCAGCAGAGATGAAGAGTCTCACTGTGACTTTGCTTGGCTGCTGATGTTCAA 720
QY 721 TACTTAGTAAATAAGCCTTCAGAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAA 780
Db 721 TACTTAGTAAATAAGCCTTCAGAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAA 780
QY 781 ATTGACGAGGAGTTTAAACAGAGCCTTGGCCAGTTGGCCTCATTTGGAATGAATTGCATT 840
```

Db 781 ATTGAGCAGGAGTTTTTAAACAGAAAGCCTTCCAGTTGGCCCTCATTGGAATGAATTGCATT 840
QY 841 TTGATGAAACAGTACATTTAGTTTGTAGCTGACAGATTACTTGTGAACTTTGGATTCTCA 900
Db 841 TTGATGAAACAGTACATTTAGTTTGTAGCTGACAGATTACTTGTGAACTTTGGATTCTCA 900
QY 901 AAGGTTTTTCAGGCAGAAATCCTTTTGTATTTTATGGAACAACTTTCTTTAGAAAGGAAA 960
Db 901 AAGGTTTTTCAGGCAGAAATCCTTTTGTATTTTATGGAACAACTTTCTTTAGAAAGGAAA 960
QY 961 ACAAAATTTCTTTGAGAAACAGATTTCAGAGTATCAGCGTTTTCAGAGTTATGGCAGAAACC 1020
Db 961 ACAAAATTTCTTTGAGAAACAGATTTCAGAGTATCAGCGTTTTCAGAGTTATGGCAGAAACC 1020
QY 1021 ACAGATAACGTTCTTACCTTTGGATGAGATT 1053
Db 1021 ACAGATAACGTTCTTACCTTTGGATGAGATT 1053

RESULT 9
US-11-088-686-18
; Sequence 18, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; PRIOR FILING DATE: 2005-03-23
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-088-686-18

Query Match 99.8%; Score 1051.4; DB 27; Length 1056;
Best Local Similarity 99.9%; Pred. No. 1.9e-270;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGGCGCCCGGAAAGCGCGGAGCGCGGCTGGATCAGATCAGAGATCATCTTCA 60
Db 1 ATGGGGCGCCCGGAAAGCGCGGAGCGCGGCTGGATCAGATCAGAGATCATCTTCA 60
QY 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
QY 121 CGGTTTGTCACTTTTCCAAATCCAGTACCTGTATTTGGAAATGTATTAACAGGCAAG 180
Db 121 CGGTTTGTCACTTTTCCAAATCCAGTACCTGTATTTGGAAATGTATTAACAGGCAAG 180
QY 181 GCTTCTCTCTGGACACGAGAGGTCGACTTATCAAGAGTCTCCTCACTGGACAAAG 240
Db 181 GCTTCTCTCTGGACACGAGAGGTCGACTTATCAAGAGTCTCCTCACTGGACAAAG 240
QY 241 CTTAAAGCAGATGAGAGTACTTCACTCTCACATCTTAGCCTTTTTCAGCGCAGTGAT 300
Db 241 CTTAAAGCAGATGAGAGTACTTCACTCTCACATCTTAGCCTTTTTCAGCGCAGTGAT 300
QY 301 GGAATTTGTAATGAAATTTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
Db 301 GGAATTTGTAATGAAATTTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
QY 361 CGCTGTTTCTATGGCTTTTCAAAATCTCATCGAGAAATGTTCACTCAGAGTGTACAGTTTG 420
Db 361 CGCTGTTTCTATGGCTTTTCAAAATCTCATCGAGAAATGTTCACTCAGAGTGTACAGTTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTTAAATGCAATTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTTTAAATGCAATTGAA 480

QY 481 ACCATGCCCTATGTTTAAAGAAAAAGCAGATTGGCCCTTGGATGATAGAGATAGAAAA 540
Db 481 ACCATGCCCTATGTTTAAAGAAAAAGCAGATTGGCCCTTGGATGATAGAGATAGAAAA 540
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCCTTTGCTGTAGAAAGGAGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCCTTTGCTGTAGAAAGGAGTTTCTTCTCAGGA 600
QY 601 TCTTTTGTCTATATCTGGCTTAAAGAAAGAGGCTTATGCCAGGACTCCTTTTTC 660
Db 601 TCTTTTGTCTATATCTGGCTTAAAGAAAGAGGCTTATGCCAGGACTCCTTTTTC 660
QY 661 AATGAATCTATCAGCAGAGATGAAGGACTTCTACCTGTGCTTTGCTGTGCTGATGTTCCAA 720
Db 661 AATGAATCTATCAGCAGAGATGAAGGACTTCTACCTGTGCTTTGCTGTGCTGATGTTCCAA 720
QY 721 TACTTAGTAAATAAGCCTTCCAGAAAGAGGCTCAGGAGAGATCATTTGTTGATGCTGTCAAA 780
Db 721 TACTTAGTAAATAAGCCTTCCAGAAAGAGGCTCAGGAGAGATCATTTGTTGATGCTGTCAAA 780
QY 781 ATTGAGCAGGAGTTTTTAAACAGAAAGCCTTGCAGTTGGCCCTCATTTGGAATGAATTGCATT 840
Db 781 ATTGAGCAGGAGTTTTTAAACAGAAAGCCTTGCAGTTGGCCCTCATTTGGAATGAATTGCATT 840
QY 841 TTGATGAAACAGTACATTTAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
Db 841 TTGATGAAACAGTACATTTAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
QY 901 AAGGTTTTTTCAGGCAGAAATCCTTTTGTATTTTATGGAACAACTTTCTTTAGAAAGGAAA 960
Db 901 AAGGTTTTTTCAGGCAGAAATCCTTTTGTATTTTATGGAACAACTTTCTTTAGAAAGGAAA 960
QY 961 ACAAAATTTCTTTGAGAAACAGTTTTCAGAGTATCAGCGTTTTCAGAGTTATGGCAGAAACC 1020
Db 961 ACAAAATTTCTTTGAGAAACAGTTTTCAGAGTATCAGCGTTTTCAGAGTTATGGCAGAAACC 1020
QY 1021 ACAGATAACGTTCTTACCTTTGGATGAGATT 1053
Db 1021 ACAGATAACGTTCTTACCTTTGGATGAGATT 1053

RESULT 10
PCT-US05-09639-20
; Sequence 20, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US05-09639-20

Query Match 99.7%; Score 1049.8; DB 1; Length 1056;
Best Local Similarity 99.8%; Pred. No. 5.1e-270;
Matches 1051; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGGGCGCCCGGAAAGCGCGGAGCGCGGCTGGATCAGATGAGATCATCTTCA 60
Db 1 ATGGGGCGCCCGGAAAGCGCGGAGCGCGGCTGGATCAGATGAGATCATCTTCA 60
QY 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGGAGTTCTCGC 120
Db 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGGAGTTCTCGC 120

```
121 CGGTTTGTCACTTTTCCAAATCCAGTACCCTGATATTTGGAAATGATATAAACAGGCGACAG 180
Db |||||
121 CGGTTTGTCACTTTTCCAAATCCAGTACCCTGATATTTGGAAATGATATAAACAGGCGACAG 180
Qy 181 GCTTCCTTCTGGACAGCAGAGAGTGCAGTATCAAAGATCTCCCTCACTGGAAACAAG 240
Db |||||
181 GCTTCCTTCTGGACAGCAGAGAGTGCAGTATCAAAGATCTCCCTCACTGGAAACAAG 240
Qy 241 CTTAAAGCAGATGAGAGTACTTCACTCTCACTCATCTTACCTTTTTCGAGCCAGTGAT 300
Db |||||
241 CTTAAAGCAGATGAGAGTACTTCACTCTCACTCATCTTACCTTTTTCGAGCCAGTGAT 300
Qy 301 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
Db |||||
301 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
Qy 361 CGCTGTTTCTATGGCTTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db |||||
361 CGCTGTTTCTATGGCTTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db |||||
421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Qy 481 ACCATGCCCTATGTTAAGAAAGAGAGATTTGGGCCCTTGCATGTGATAGCAGATAGAAA 540
Db |||||
481 ACCATGCCCTATGTTAAGAAAGAGAGATTTGGGCCCTTGCATGTGATAGCAGATAGAAA 540
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGAGTTCCTCTCAGGA 600
Db |||||
541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGAGTTCCTCTCAGGA 600
Qy 601 TCTTTTGTCTGCTATATCTGGCTTAAGAGAGAGAGTCTTATGCGCAGACTCACTTTTTC 660
Db |||||
601 TCTTTTGTCTGCTATATCTGGCTTAAGAGAGAGAGTCTTATGCGCAGACTCACTTTTTC 660
Qy 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTGCTGATGTTCCAA 720
Db |||||
661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTGCTGATGTTCCAA 720
Qy 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGTGCAGGATCATTTGTGTGATGTCTGTCAAA 780
Db |||||
721 TACTTAGTAAATAAGCCTTCAGAAAGAGGTGCAGGATCATTTGTGTGATGTCTGTCAAA 780
Qy 781 ATTGACGAGAGTTTTTAACAGAGCCTTCCAGTTGGCCTCATTTGGAATGAAATTGCATT 840
Db |||||
781 ATTGACGAGAGTTTTTAACAGAGCCTTCCAGTTGGCCTCATTTGGAATGAAATTGCATT 840
Qy 841 TTGATGAACAGTACATTGAGTTTGTAGCTGCACAGATTACTTGTGGAACTTGGATTCTCA 900
Db |||||
841 TTGATGAACAGTACATTGAGTTTGTAGCTGCACAGATTACTTGTGGAACTTGGATTCTCA 900
Qy 901 AAGGTTTTTCAGCAGAAAATCCTTTGATTTTATGGAAACAAATTTCTTTAGAGGAAA 960
Db |||||
901 AAGGTTTTTCAGCAGAAAATCCTTTGATTTTATGGAAACAAATTTCTTTAGAGGAAA 960
Qy 961 ACAATTTCTTTGAGAAAACAGTTTCAGAGTATCAGGTTTTCAGTTATGGCAGAAACC 1020
Db |||||
961 ACAATTTCTTTGAGAAAACAGTTTCAGAGTATCAGGTTTTCAGTTATGGCAGAAACC 1020
Qy 1021 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1053
Db |||||
1021 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1053
```

RESULT 11

US-11-088-686-20
; Sequence 20, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686

; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-088-686-20

Query Match 99.7%; Score 1049.8; DB 27; Length 1056;
Best Local Similarity 99.8%; Pred. No. 5.1e-270;
Matches 1051; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy 1 ATGGCGACCCCGGAAAGCGCGGCGGCTGATCAGGATGAGATCATCTTCA 60
Db 1 ATGGCGACCCCGGAAAGCGCGGCGGCTGATCAGGATGAGATCATCTTCA 60
Qy 61 GACACCAACGAAAGTGAATATAAGTCAAAATGAAGAGCCACTCCCTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAACGAAAGTGAATATAAGTCAAAATGAAGAGCCACTCCCTAAGAAAGAGTTCTCGC 120
Qy 121 CGGTTTGTCACTTTTCCAAATCCAGTACCCTGATATTTGGAAATGATATAAACAGGCGACAG 180
Db 121 CGGTTTGTCACTTTTCCAAATCCAGTACCCTGATATTTGGAAATGATATAAACAGGCGACAG 180
Qy 181 GCTTCCTTCTGGACAGCAGAGAGTGCAGTATCAAAGATCTCCCTCACTGGAACAG 240
Db 181 GCTTCCTTCTGGACAGCAGAGAGTGCAGTATCAAAGATCTCCCTCACTGGAACAG 240
Qy 241 CTTAAAGCAGATCAGAGTACTTCACTCTCACTCATCTTACCTTTTTCGAGCCAGTGAT 300
Db 241 CTTAAAGCAGATCAGAGTACTTCACTCTCACTCATCTTACCTTTTTCGAGCCAGTGAT 300
Qy 301 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
Db 301 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 360
Qy 361 CGCTGTTTCTATGGCTTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGGCTTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Qy 481 ACCATGCCCTATGTTAAGAAAGAGAGATTTGGGCCCTTGCATGTGATAGCAGATAGAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAGAGAGATTTGGGCCCTTGCATGTGATAGCAGATAGAAA 540
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGTTCCTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGTTCCTCTCAGGA 600
Qy 601 TCTTTTGTCTGCTATATCTGGCTTAAGAGAGAGTCTTATGCGCAGACTCACTTTTTC 660
Db 601 TCTTTTGTCTGCTATATCTGGCTTAAGAGAGAGTCTTATGCGCAGACTCACTTTTTC 660
Qy 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTGCTGATGTTCCAA 720
Db 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTGCTGATGTTCCAA 720
Qy 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGTGCAGGATCATTTGTGTGATGTCTGTCAAA 780
Db 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGTGCAGGATCATTTGTGTGATGTCTGTCAAA 780
Qy 781 ATTGACGAGAGTTTTTAACAGAGCCTTCCAGTTGGCCTCATTTGGAATGAAATTGCATT 840
Db 781 ATTGACGAGAGTTTTTAACAGAGCCTTCCAGTTGGCCTCATTTGGAATGAAATTGCATT 840
Qy 841 TTGATGAACAGTACATTGAGTTTGTAGCTGCACAGATTACTTGTGGAACTTGGATTCTCA 900
Db |||||
```

Db 841 TTGATGAAACAGTACATGAGTTGTAGCTGACAGATTAATCTGTGAACTTGGATTTCTCA 900
QY 901 AAGGTTTTTTCAGGCAGAAAATCCTTTTGTATTTTATGAAAAACATTTCTTTAGAAAGAAA 960
Db 901 AAGGTTTTTTCAGGCAGAAAATCCTTTTGTATTTTATGAAAAACATTTCTTTAGAAAGAAA 960
QY 961 ACAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTTGCAGTTATGCGCAGAAACC 1020
Db 961 ACAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTTGCAGTTATGCGCAGAAACC 1020
QY 1021 ACAGATAACGCTCTTACCTTGGATGAGATTTT 1053
Db 1021 ACAGATAACGCTCTTACCTTGGATGAGATTTT 1053

RESULT 12

PCT-US05-09639-22
; Sequence 22, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US05-09639-22

Query Match 99.5%; Score 1048.2; DB 1; Length 1056;
Best Local Similarity 99.7%; Pred. No. 1.4e-269;
Matches 1050; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAGCGCGGAGCGCGGGCTGGATCAGATGAGATCACTTTCA 60
Db 1 ATGGGCGACCCGGAAGCGCGGAGCGCGGGCTGGATCAGATGAGATCACTTTCA 60
QY 61 GACACCAACGAAAGTGAATAAGTCAATGAAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAACGAAAGTGAATAAGTCAATGAAGCCACTCTTAAGAAAGAGTTCTCGC 120
QY 121 CGTTTGTGTCATCTTTCCAAATCCAGTACCTGATATTTGAAAATGTATAACAGGCACAG 180
Db 121 CGTTTGTGTCATCTTTCCAAATCCAGTACCTGATATTTGAAAATGTATAACAGGCACAG 180
QY 181 GCTTCCTCTTGACACGACGAGAGTGCATCTTATCAAGAGTCTCCCTCACTGGAACAG 240
Db 181 GCTTCCTCTTGACACGACGAGAGTGCATCTTATCAAGAGTCTCCCTCACTGGAACAG 240
QY 241 CTTTAAAGCAGATGAGAAGTACTCTCTCTCATCTTAGCCCTTTTTCGACCGAGTGT 300
Db 241 CTTTAAAGCAGATGAGAAGTACTCTCTCTCATCTTAGCCCTTTTTCGACCGAGTGT 300
QY 301 GGAATTTGTAATCAAAATTTGTGGAGCGCTTTAGTCAGAGGTGCGAGTTCCAGAGGCT 360
Db 301 GGAATTTGTAATCAAAATTTGTGGAGCGCTTTAGTCAGAGGTGCGAGTTCCAGAGGCT 360
QY 361 CGTGTGTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGTGTGTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
QY 421 CTGATAGACATTAATCAGAGATCCCAAGAAAGGAAATTTTAAATCAATTTGA 480
Db 421 CTGATAGACATTAATCAGAGATCCCAAGAAAGGAAATTTTAAATCAATTTGA 480
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCTTCGGATGGATAGCAGATAGAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCTTCGGATGGATAGCAGATAGAAA 540

QY 541 TCTACTTTTGGGAAAAAGAGTGTGGCTTTGCTGCTGTAGAGAGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAAAGAGTGTGGCTTTGCTGCTGTAGAGAGTTTCTTCTCAGGA 600
QY 601 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACATTTTCC 660
Db 601 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCACATTTTCC 660
QY 661 AATGAATCTCATCAGCAGAGATGAGGACTTCACTGTGACTTTTGTGCTGTGCTGATGTTCCAA 720
Db 661 AATGAATCTCATCAGCAGAGATGAGGACTTCACTGTGACTTTTGTGCTGTGCTGATGTTCCAA 720
QY 721 TACTTAGTAAATAAGCCTTTCAGAAAGAGGTTCAGGAGATCAATTTGTGATGCTGTCAAA 780
Db 721 TACTTAGTAAATAAGCCTTTCAGAAAGAGGTTCAGGAGATCAATTTGTGATGCTGTCAAA 780
QY 781 ATTGAGCAGGAGTTTTTAAACAGAAAGCCTTTCAGAAAGGTTCAGGAGATCAATTTGTGATGCTGTCA 840
Db 781 ATTGAGCAGGAGTTTTTAAACAGAAAGCCTTTCAGAAAGGTTCAGGAGATCAATTTGTGATGCTGTCA 840
QY 841 TTGATGAAACAGTACATTTGAGTTTGTAGCTGACAGATTAATCTTGTGAACTTGGATTTCTCA 900
Db 841 TTGATGAAACAGTACATTTGAGTTTGTAGCTGACAGATTAATCTTGTGAACTTGGATTTCTCA 900
QY 901 AAGGTTTTTTCAGGCAGAAAATCCTTTTGTATTTTATGAAAACATTTCTTTAGAAAGAAA 960
Db 901 AAGGTTTTTTCAGGCAGAAAATCCTTTTGTATTTTATGAAAACATTTCTTTAGAAAGAAA 960
QY 961 ACAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTTGCAGTTATGCGCAGAAACC 1020
Db 961 ACAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTTGCAGTTATGCGCAGAAACC 1020
QY 1021 ACAGATAACGCTCTTACCTTGGATGAGATTTT 1053
Db 1021 ACAGATAACGCTCTTACCTTGGATGAGATTTT 1053

RESULT 13

PCT-US05-09639-24
; Sequence 24, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US05-09639-24

Query Match 99.5%; Score 1048.2; DB 1; Length 1056;
Best Local Similarity 99.7%; Pred. No. 1.4e-269;
Matches 1050; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAGCGCGGAGCGCGGGCTGGATCAGATGAGATCACTTTCA 60
Db 1 ATGGGCGACCCGGAAGCGCGGAGCGCGGGCTGGATCAGATGAGATCACTTTCA 60
QY 61 GACACCAACGAAAGTGAATAAGTCAATGAAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 61 GACACCAACGAAAGTGAATAAGTCAATGAAGCCACTCTTAAGAAAGAGTTCTCGC 120
QY 121 CGTTTGTGTCATCTTTCCAAATCCAGTACCTGATATTTGAAAATGTATAACAGGCACAG 180
Db 121 CGTTTGTGTCATCTTTCCAAATCCAGTACCTGATATTTGAAAATGTATAACAGGCACAG 180


```
181 GCTTCTCTCTGACAGCAGAGAGGTCGACTTATCAAAGGATCTCCCTCACTGGAAACAG 240
181 GCTTCTCTCTGACAGCAGAGAGGTCGACTTATCAAAGGATCTCCCTCACTGGAAACAG 240
241 CTTAAAGCAGATGAGAAAGTACTTCTCATCTCTCACTCTTAAAGGATCTCCCTCACTGGAAACAG 300
241 CTTAAAGCAGATGAGAAAGTACTTCTCATCTCTCACTCTTAAAGGATCTCCCTCACTGGAAACAG 300
301 GGAATTTGTAATGAAATTTGTTGGAGCGCTTTAGTCAGGAGGTCAGGATCCAGAGGCT 360
301 GGAATTTGTAATGAAATTTGTTGGAGCGCTTTAGTCAGGAGGTCAGGATCCAGAGGCT 360
361 CGCTGTTCTATGCGCTTTCAAAATCTCATCGAGATGTTCACTCAGAGATGTCAGATTTG 420
361 CGCTGTTCTATGCGCTTTCAAAATCTCATCGAGATGTTCACTCAGAGATGTCAGATTTG 420
421 CTGATAGACACTTATCAGAGATCCCAAGAAAGGGAATTTTAAATGCAATTTGAA 480
421 CTGATAGACACTTATCAGAGATCCCAAGAAAGGGAATTTTAAATGCAATTTGAA 480
481 ACCATGCGCTATGTTAAAGAAAGCAGATTTGGGCTTCGCGATGGATAGCAGATAGAAA 540
481 ACCATGCGCTATGTTAAAGAAAGCAGATTTGGGCTTCGCGATGGATAGCAGATAGAAA 540
541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAGAGGATTTTCTCTCAGGA 600
541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAGAGGATTTTCTCTCAGGA 600
601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGTCCTTATGCCAGGACTCACTTTTCC 660
601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGTCCTTATGCCAGGACTCACTTTTCC 660
661 AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTGTGCTGCTGATGTCGAA 720
661 AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTGTGCTGCTGATGTCGAA 720
721 TACTTAGTAAATGAGCTTTCAGAGAAAGGTCAGGAGATCAATTTGATGCTGTCAAA 780
721 TACTTAGTAAATGAGCTTTCAGAGAAAGGTCAGGAGATCAATTTGATGCTGTCAAA 780
781 ATTGAGCAGGAGTCTTAAACAGAGCCTTGGCCCTCAATGGAAATGAATTCGATT 840
781 ATTGAGCAGGAGTCTTAAACAGAGCCTTGGCCCTCAATGGAAATGAATTCGATT 840
841 TTGATGAAACAGTACATTCAGTCTTGTAGCTGACAGATTAATTCGTAATTCGATTCCTCA 900
841 TTGATGAAACAGTACATTCAGTCTTGTAGCTGACAGATTAATTCGTAATTCGATTCCTCA 900
901 AAGGTTTTTTCAGGCAGAAATCCTTTTGAATTTTATGGAACATTTCTTTAGAGGAAAC 960
901 AAGGTTTTTTCAGGCAGAAATCCTTTTGAATTTTATGGAACATTTCTTTAGAGGAAAC 960
961 ACAAAATTTCTTTGAGAAACAGTCTTTCAGAGTATCAGCGTTTTCAGATTTGCGAGTAAC 1020
961 ACAAAATTTCTTTGAGAAACAGTCTTTCAGAGTATCAGCGTTTTCAGATTTGCGAGTAAC 1020
1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053
1021 ACAGATAACGCTTTCACCTTGGATGAGATTTT 1053
```

RESULT 14

```
US-11-088-686-22
; Sequence 22, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-088-686-22

Query Match          99.5%; Score 1048.2; DB 27; Length 1056;
Best Local Similarity 99.7%; Pred. No. 1.4e-269; Indels 0; Gaps 0;
Matches 1050; Conservative 0; Mismatches 3;

QY 1 ATGGGCGACCCGGAAGCGCGGCTGGATCAGGATCAGAGATCATCTTCA 60
DB 1 ATGGGCGACCCGGAAGCGCGGCTGGATCAGGATCAGAGATCATCTTCA 60
QY 61 GACACCAAGCAAGTGAATTAAGTCAAAATGAAGAGCCTCTCTAAGAAAGATTTCTCGC 120
DB 61 GACACCAAGCAAGTGAATTAAGTCAAAATGAAGAGCCTCTCTAAGAAAGATTTCTCGC 120
QY 121 CGGTTTGTCTATCTTTCCAAATCCAGTACCCTTGATATTTGGAAATGTATAACAGGCACAG 180
DB 121 CGGTTTGTCTATCTTTCCAAATCCAGTACCCTTGATATTTGGAAATGTATAACAGGCACAG 180
QY 181 GCTTCTCTCTGACAGCAGAGAGGTCGACTTATCAAAGGATCTCCCTCACTGGAAACAG 240
DB 181 GCTTCTCTCTGACAGCAGAGAGGTCGACTTATCAAAGGATCTCCCTCACTGGAAACAG 240
QY 241 CTTAAAGCAGATGAGAAAGTACTTCTCATCTCTCACTCTTAGCCTTTTTCGAGCAGTGAT 300
DB 241 CTTAAAGCAGATGAGAAAGTACTTCTCATCTCTCACTCTTAGCCTTTTTCGAGCAGTGAT 300
QY 301 GGAATTTGTAATGAAATTTGTTGGAGCGCTTTAGTCAGGAGGTCAGGTTCCAGAGGCT 360
DB 301 GGAATTTGTAATGAAATTTGTTGGAGCGCTTTAGTCAGGAGGTCAGGTTCCAGAGGCT 360
QY 361 CGCTGTTCTATGCGCTTTCAAAATCTCATCGAGATGTTCACTCAGAGATGTCAGATTTG 420
DB 361 CGCTGTTCTATGCGCTTTCAAAATCTCATCGAGATGTTCACTCAGAGATGTCAGATTTG 420
QY 421 CTGATAGACACTTATCAGAGATCCCAAGAAAGGGAATTTTAAATGCAATTTGAA 480
DB 421 CTGATAGACACTTATCAGAGATCCCAAGAAAGGGAATTTTAAATGCAATTTGAA 480
QY 481 ACCATGCGCTATGTTAAAGAAAGCAGATTTGGGCTTCGCGATGGATAGCAGATAGAAA 540
DB 481 ACCATGCGCTATGTTAAAGAAAGCAGATTTGGGCTTCGCGATGGATAGCAGATAGAAA 540
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAGAGGATTTTCTCTCAGGA 600
DB 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAGAGGATTTTCTCTCAGGA 600
QY 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGTCCTTATGCCAGGACTCACTTTTCC 660
DB 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGTCCTTATGCCAGGACTCACTTTTCC 660
QY 661 AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTGTGCTGCTGATGTCGAA 720
DB 661 AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTGTGCTGCTGATGTCGAA 720
QY 721 TACTTAGTAAATGAGCTTTCAGAGAAAGGTCAGGAGATCAATTTGATGCTGTCAAA 780
DB 721 TACTTAGTAAATGAGCTTTCAGAGAAAGGTCAGGAGATCAATTTGATGCTGTCAAA 780
QY 781 ATTGAGCAGGAGTCTTAAACAGAGCCTTGGCCCTCAATGGAAATGAATTCGATT 840
DB 781 ATTGAGCAGGAGTCTTAAACAGAGCCTTGGCCCTCAATGGAAATGAATTCGATT 840
QY 841 TTGATGAAACAGTACATTCAGTCTTGTAGCTGACAGATTAATTCGTAATTCGATTCCTCA 900
DB 841 TTGATGAAACAGTACATTCAGTCTTGTAGCTGACAGATTAATTCGTAATTCGATTCCTCA 900
QY 901 AAGGTTTTTTCAGGCAGAAATCCTTTTGAATTTTATGGAACATTTCTTTAGAGGAAAC 960
DB 901 AAGGTTTTTTCAGGCAGAAATCCTTTTGAATTTTATGGAACATTTCTTTAGAGGAAAC 960
```



```
Db 901 AAGTTTTTCAGCAGAGAAATCCTTTTGATTTTATGGAAGAAACATTTCTTTAGAGAGAAA 960
Qy 961 ACAATTTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTCAGTTCATGCGAGAAACC 1020
Db 961 ACAATTTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTCAGTTCATGCGAGAAACC 1020
Qy 1021 ACAGATAACGTCTTTCACCTTGGATGAGATTTT 1053
Db 1021 ACAGATAACGTCTTTCACCTTGGATGAGATTTT 1053

RESULT 15
US-11-088-686-24
; Sequence 24, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; PRIOR FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-088-686-24

Query Match 99.5%; Score 1048.2; DB 27; Length 1056;
Best Local Similarity 99.7%; Pred. No. 1.4e-269;
Matches 1050; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATGGCGACCCGGAAGCGCGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 1 ATGGCGACCCGGAAGCGCGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Qy 61 GACCCNACGAAAGTGAATTAAGTCAATGAAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 61 GACCCNACGAAAGTGAATTAAGTCAATGAAGCCACTCTTAAGAAAGAGTTCTCGC 120
Qy 121 CGGTTTGTCTATCTTCCATCCAGTACCCTGATATTTTGAAGATGATATAACAGGCAAG 180
Db 121 CGGTTTGTCTATCTTCCATCCAGTACCCTGATATTTTGAAGATGATATAACAGGCAAG 180
Qy 181 GCTTCCTCTGGACAGCAGAGAGTGCTGATATCAAGGATCTCCCTCACTGGAACAAG 240
Db 181 GCTTCCTCTGGACAGCAGAGAGTGCTGATATCAAGGATCTCCCTCACTGGAACAAG 240
Qy 241 CTTAAGCAGATGAGAAAGTACTTCTCATCTTACATCTTACCTTTTTCAGCCAGTAT 300
Db 241 CTTAAGCAGATGAGAAAGTACTTCTCATCTTACATCTTACCTTTTTCAGCCAGTAT 300
Qy 301 GGAATTTGTAATGAAATTTTGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
Db 301 GGAATTTGTAATGAAATTTTGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
Qy 361 CGCTGTTTCTATGGCTTTTCAAAATTTCTATCGAGAATGTTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTTCTATGGCTTTTCAAAATTTCTATCGAGAATGTTTCACTCAGAGATGTACAGTTTG 420
Qy 421 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Db 421 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
Qy 481 ACCATGCCCTATGTTAAGAAAGAAAGCAGATTCGGCTTCGGATGGATGAGATAGAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAGAAAGCAGATTCGGCTTCGGATGGATGAGATAGAAA 540
Qy 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGTGCTGTAGAGAGGATTTTCTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGTGCTGTAGAGAGGATTTTCTCTCAGGA 600
```

```
Qy 601 TCTTTTCTGTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCCTTTTTC 660
Db 601 TCTTTTCTGTATATTTCTGGCTAAAGAGAGAGGTCTTATGCCAGGACTCCTTTTTC 660
Qy 661 AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGTGCTGATGTTCCAA 720
Db 661 AATGAACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGTGCTGATGTTCCAA 720
Qy 721 TACTTAGTAATAAGCCCTTCAGAGAAAGGCTCAGGAGATCATTTGATGCTCTCAAA 780
Db 721 TACTTAGTAATAAGCCCTTCAGAGAAAGGCTCAGGAGATCATTTGATGCTCTCAAA 780
Qy 781 ATTGAGCAGGAGTTTTTAAACAGAGCCCTTGGCAGTTGGCCCTCATTTGGAATGAATTCATT 840
Db 781 ATTGAGCAGGAGTTTTTAAACAGAGCCCTTGGCAGTTGGCCCTCATTTGGAATGAATTCATT 840
Qy 841 TTGATGAAACAGTACATTGAGTTTTGTAGCTGACAGATTTACTTTGTGAACTTTGGATTCTCA 900
Db 841 TTGATGAAACAGTACATTGAGTTTTGTAGCTGACAGATTTACTTTGTGAACTTTGGATTCTCA 900
Qy 901 AAGTTTTTTCAGGAGAGAAATCCTTTTGAATTTTATGAAACATTTCTTTAGAGAGAAA 960
Db 901 AAGTTTTTTCAGGAGAGAAATCCTTTTGAATTTTATGAAACATTTCTTTAGAGAGAAA 960
Qy 961 ACAATTTTCTTTGAGAAACAGATTTTCAAGATATCAGCGTTTTTGCAGTTTATGSCAGAAACC 1020
Db 961 ACAATTTTCTTTGAGAAACAGATTTTCAAGATATCAGCGTTTTTGCAGTTTATGSCAGAAACC 1020
Qy 1021 ACAGATAACGTCTTTCACCTTGGATGAGATTTT 1053
Db 1021 ACAGATAACGTCTTTCACCTTGGATGAGATTTT 1053
```

Search completed: October 30, 2005, 06:01:25
Job time : 1375.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2005, 23:48:01 ; Search time 3214.5 Seconds
(without alignments)
13372.738 Million cell updates/sec

Title: US-10-698-228-2

Perfect score: 1053

Sequence: 1 atggcgacccgaaggcc.....tcacctggatgcagatttt 1053

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 45554873 seqs, 20411521753 residues

Total number of hits satisfying chosen parameters: 91109746

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents NA Main:*

- 1: /cgn2_6/ptodata/1/pna/PCTUS1 COMB.seq.*
- 2: /cgn2_6/ptodata/1/pna/PCTUS2 COMB.seq.*
- 3: /cgn2_6/ptodata/1/pna/PCTUS COMB.seq.*
- 4: /cgn2_6/ptodata/1/pna/US06 COMB.seq.*
- 5: /cgn2_6/ptodata/1/pna/US07 COMB.seq.*
- 6: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
- 7: /cgn2_6/ptodata/1/pna/US081 COMB.seq.*
- 8: /cgn2_6/ptodata/1/pna/US082 COMB.seq.*
- 9: /cgn2_6/ptodata/1/pna/US083 COMB.seq.*
- 10: /cgn2_6/ptodata/1/pna/US084 COMB.seq.*
- 11: /cgn2_6/ptodata/1/pna/US085 COMB.seq.*
- 12: /cgn2_6/ptodata/1/pna/US086 COMB.seq.*
- 13: /cgn2_6/ptodata/1/pna/US087 COMB.seq.*
- 14: /cgn2_6/ptodata/1/pna/US088 COMB.seq.*
- 15: /cgn2_6/ptodata/1/pna/US089 COMB.seq.*
- 16: /cgn2_6/ptodata/1/pna/US090 COMB.seq.*
- 17: /cgn2_6/ptodata/1/pna/US091 COMB.seq.*
- 18: /cgn2_6/ptodata/1/pna/US092A COMB.seq.*
- 19: /cgn2_6/ptodata/1/pna/US092B COMB.seq.*
- 20: /cgn2_6/ptodata/1/pna/US093A COMB.seq.*
- 21: /cgn2_6/ptodata/1/pna/US093B COMB.seq.*
- 22: /cgn2_6/ptodata/1/pna/US094 COMB.seq.*
- 23: /cgn2_6/ptodata/1/pna/US095A COMB.seq.*
- 24: /cgn2_6/ptodata/1/pna/US095B COMB.seq.*
- 25: /cgn2_6/ptodata/1/pna/US095C COMB.seq.*
- 26: /cgn2_6/ptodata/1/pna/US095D COMB.seq.*
- 27: /cgn2_6/ptodata/1/pna/US096A COMB.seq.*
- 28: /cgn2_6/ptodata/1/pna/US096B COMB.seq.*
- 29: /cgn2_6/ptodata/1/pna/US096C COMB.seq.*
- 30: /cgn2_6/ptodata/1/pna/US096D COMB.seq.*
- 31: /cgn2_6/ptodata/1/pna/US096E COMB.seq.*
- 32: /cgn2_6/ptodata/1/pna/US097A COMB.seq.*
- 33: /cgn2_6/ptodata/1/pna/US097B COMB.seq.*
- 34: /cgn2_6/ptodata/1/pna/US097C COMB.seq.*
- 35: /cgn2_6/ptodata/1/pna/US098A COMB.seq.*
- 36: /cgn2_6/ptodata/1/pna/US098B COMB.seq.*
- 37: /cgn2_6/ptodata/1/pna/US098C COMB.seq.*
- 38: /cgn2_6/ptodata/1/pna/US098D COMB.seq.*
- 39: /cgn2_6/ptodata/1/pna/US099A COMB.seq.*
- 40: /cgn2_6/ptodata/1/pna/US099B COMB.seq.*
- 41: /cgn2_6/ptodata/1/pna/US099C COMB.seq.*
- 42: /cgn2_6/ptodata/1/pna/US099D COMB.seq.*
- 43: /cgn2_6/ptodata/1/pna/US099E COMB.seq.*

- 44: /cgn2_6/ptodata/1/pna/US099F COMB.seq.*
- 45: /cgn2_6/ptodata/1/pna/US099G COMB.seq.*
- 46: /cgn2_6/ptodata/1/pna/US100A COMB.seq.*
- 47: /cgn2_6/ptodata/1/pna/US100B COMB.seq.*
- 48: /cgn2_6/ptodata/1/pna/US101A COMB.seq.*
- 49: /cgn2_6/ptodata/1/pna/US101B COMB.seq.*
- 50: /cgn2_6/ptodata/1/pna/US102A COMB.seq.*
- 51: /cgn2_6/ptodata/1/pna/US102B COMB.seq.*
- 52: /cgn2_6/ptodata/1/pna/US103A COMB.seq.*
- 53: /cgn2_6/ptodata/1/pna/US103B COMB.seq.*
- 54: /cgn2_6/ptodata/1/pna/US104A COMB.seq.*
- 55: /cgn2_6/ptodata/1/pna/US104B COMB.seq.*
- 56: /cgn2_6/ptodata/1/pna/US105A COMB.seq.*
- 57: /cgn2_6/ptodata/1/pna/US105B COMB.seq.*
- 58: /cgn2_6/ptodata/1/pna/US106A COMB.seq.*
- 59: /cgn2_6/ptodata/1/pna/US107A COMB.seq.*
- 60: /cgn2_6/ptodata/1/pna/US107B COMB.seq.*
- 61: /cgn2_6/ptodata/1/pna/US107C COMB.seq.*
- 62: /cgn2_6/ptodata/1/pna/US107D COMB.seq.*
- 63: /cgn2_6/ptodata/1/pna/US108A COMB.seq.*
- 64: /cgn2_6/ptodata/1/pna/US108B COMB.seq.*
- 65: /cgn2_6/ptodata/1/pna/US109A COMB.seq.*
- 66: /cgn2_6/ptodata/1/pna/US109B COMB.seq.*
- 67: /cgn2_6/ptodata/1/pna/US109C COMB.seq.*
- 68: /cgn2_6/ptodata/1/pna/US110 COMB.seq.*
- 69: /cgn2_6/ptodata/1/pna/US6000 COMB.seq.*
- 70: /cgn2_6/ptodata/1/pna/US6001 COMB.seq.*
- 71: /cgn2_6/ptodata/1/pna/US6002 COMB.seq.*
- 72: /cgn2_6/ptodata/1/pna/US6003 COMB.seq.*
- 73: /cgn2_6/ptodata/1/pna/US6004 COMB.seq.*
- 74: /cgn2_6/ptodata/1/pna/US6005 COMB.seq.*
- 75: /cgn2_6/ptodata/1/pna/US6006 COMB.seq.*
- 76: /cgn2_6/ptodata/1/pna/US6007 COMB.seq.*
- 77: /cgn2_6/ptodata/1/pna/US6008 COMB.seq.*
- 78: /cgn2_6/ptodata/1/pna/US6009 COMB.seq.*
- 79: /cgn2_6/ptodata/1/pna/US6010 COMB.seq.*
- 80: /cgn2_6/ptodata/1/pna/US6011 COMB.seq.*
- 81: /cgn2_6/ptodata/1/pna/US6012 COMB.seq.*
- 82: /cgn2_6/ptodata/1/pna/US6013 COMB.seq.*
- 83: /cgn2_6/ptodata/1/pna/US6014 COMB.seq.*
- 84: /cgn2_6/ptodata/1/pna/US6015 COMB.seq.*
- 85: /cgn2_6/ptodata/1/pna/US6016 COMB.seq.*
- 86: /cgn2_6/ptodata/1/pna/US6017 COMB.seq.*
- 87: /cgn2_6/ptodata/1/pna/US6018 COMB.seq.*
- 88: /cgn2_6/ptodata/1/pna/US6019 COMB.seq.*
- 89: /cgn2_6/ptodata/1/pna/US6020 COMB.seq.*
- 90: /cgn2_6/ptodata/1/pna/US6021 COMB.seq.*
- 91: /cgn2_6/ptodata/1/pna/US6022 COMB.seq.*
- 92: /cgn2_6/ptodata/1/pna/US6023A COMB.seq.*
- 93: /cgn2_6/ptodata/1/pna/US6023B COMB.seq.*
- 94: /cgn2_6/ptodata/1/pna/US6024 COMB.seq.*
- 95: /cgn2_6/ptodata/1/pna/US6025 COMB.seq.*
- 96: /cgn2_6/ptodata/1/pna/US6026 COMB.seq.*
- 97: /cgn2_6/ptodata/1/pna/US6027 COMB.seq.*
- 98: /cgn2_6/ptodata/1/pna/US6028 COMB.seq.*
- 99: /cgn2_6/ptodata/1/pna/US6029 COMB.seq.*
- 100: /cgn2_6/ptodata/1/pna/US6030 COMB.seq.*
- 101: /cgn2_6/ptodata/1/pna/US6031 COMB.seq.*
- 102: /cgn2_6/ptodata/1/pna/US6032 COMB.seq.*
- 103: /cgn2_6/ptodata/1/pna/US6033 COMB.seq.*
- 104: /cgn2_6/ptodata/1/pna/US6034 COMB.seq.*
- 105: /cgn2_6/ptodata/1/pna/US6035 COMB.seq.*
- 106: /cgn2_6/ptodata/1/pna/US6036 COMB.seq.*
- 107: /cgn2_6/ptodata/1/pna/US6037 COMB.seq.*
- 108: /cgn2_6/ptodata/1/pna/US6038 COMB.seq.*
- 109: /cgn2_6/ptodata/1/pna/US6039 COMB.seq.*
- 110: /cgn2_6/ptodata/1/pna/US6040 COMB.seq.*
- 111: /cgn2_6/ptodata/1/pna/US6041 COMB.seq.*
- 112: /cgn2_6/ptodata/1/pna/US6042 COMB.seq.*
- 113: /cgn2_6/ptodata/1/pna/US6043 COMB.seq.*
- 114: /cgn2_6/ptodata/1/pna/US6044 COMB.seq.*
- 115: /cgn2_6/ptodata/1/pna/US6045 COMB.seq.*
- 116: /cgn2_6/ptodata/1/pna/US6046 COMB.seq.*

117: /cgn2_6/ptodata/1/pna/US6047 COMB.seq.*
118: /cgn2_6/ptodata/1/pna/US6048 COMB.seq.*
119: /cgn2_6/ptodata/1/pna/US6049 COMB.seq.*
120: /cgn2_6/ptodata/1/pna/US6050 COMB.seq.*
121: /cgn2_6/ptodata/1/pna/US6051 COMB.seq.*
122: /cgn2_6/ptodata/1/pna/US6052 COMB.seq.*
123: /cgn2_6/ptodata/1/pna/US6053 COMB.seq.*
124: /cgn2_6/ptodata/1/pna/US6054 COMB.seq.*
125: /cgn2_6/ptodata/1/pna/US6055 COMB.seq.*
126: /cgn2_6/ptodata/1/pna/US6056 COMB.seq.*
127: /cgn2_6/ptodata/1/pna/US6057 COMB.seq.*
128: /cgn2_6/ptodata/1/pna/US6058 COMB.seq.*
129: /cgn2_6/ptodata/1/pna/US6059 COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1053	100.0	1053	58	US-10-698-228-2
2	1053	100.0	2596	34	Sequence 28, Appli
3	1053	100.0	2596	81	Sequence 20, Appli
4	1053	100.0	3397	49	Sequence 20856, A
5	1053	100.0	4955	1	Sequence 71, Appli
6	1053	100.0	4955	2	Sequence 71, Appli
7	1053	100.0	4955	49	Sequence 71, Appli
8	1053	100.0	4955	52	Sequence 71, Appli
9	1053	100.0	4955	58	Sequence 3, Appli
10	1052.6	100.0	1849	67	Sequence 49, Appli
11	1052.6	100.0	1849	120	Sequence 20, Appli
12	1052.6	100.0	3397	114	Sequence 12, Appli
13	1051.4	99.8	1053	58	Sequence 12, Appli
14	1051.4	99.8	1081	58	Sequence 4, Appli
15	1049.8	99.7	1601	28	Sequence 12810, A
16	1049.8	99.7	1601	65	Sequence 12810, A
17	1042	99.0	2597	97	Sequence 266, App
18	876.8	83.3	5609	26	Sequence 6398, Ap
19	876.8	83.3	1140	48	Sequence 870, App
20	876.8	83.3	1140	106	Sequence 870, App
21	650.6	61.8	706	2	Sequence 8394, Ap
22	650.6	61.8	706	56	Sequence 8394, Ap
23	650.6	61.8	706	107	Sequence 8394, Ap
24	579	55.0	2500	29	Sequence 9, Appli
25	579	55.0	2500	29	Sequence 9, Appli
26	578.4	54.9	1649	2	Sequence 583, App
27	578.4	54.9	1989	1	Sequence 505, App
28	578.4	54.9	1989	40	Sequence 505, App
29	578.4	54.9	2215	2	Sequence 505, App
30	578.4	54.9	2216	47	Sequence 38, Appli
31	578.4	54.9	2216	47	Sequence 342, App
32	578.4	54.9	2216	86	Sequence 422, App
33	578.4	54.9	2479	66	Sequence 2025, Ap
34	578.4	54.9	2482	1	Sequence 514, App
35	578.4	54.9	2482	23	Sequence 514, App
36	578.4	54.9	2482	26	Sequence 13666, A
37	578.4	54.9	2482	51	Sequence 13666, A
38	578.4	54.9	2500	1	Sequence 514, App
39	578.4	54.9	2500	42	Sequence 4, Appli
40	578.4	54.9	2500	42	Sequence 724, App
41	578.4	54.9	2500	42	Sequence 1169, Ap
42	578.4	54.9	2500	42	Sequence 1827, Ap
43	578.4	54.9	2500	58	Sequence 1370, Ap
44	578.4	54.9	2500	62	Sequence 458, App
45	578.4	54.9	2500	62	Sequence 713, App
					Sequence 3751, Ap

ALIGNMENTS

RESULT 1
US-10-698-228-2
; Sequence 2, Application US/10698228
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-2

Query Match 100.0%; Score 1053; DB 58; Length 1053;
Best Local Similarity 100.0%; Pred. No. 2e-276;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGCAGCCCGGAAAGCGCGGCGGCTGGATCAGATGAGATCATCTTCA 60
Db 1 ATGGCGCAGCCCGGAAAGCGCGGCGGCTGGATCAGATGAGATCATCTTCA 60

QY 61 GACACCAACGAAAGTGAATATAAGTCAATGAAGAGCCACTCTTCAAGAAAGTTCGC 120
Db 61 GACACCAACGAAAGTGAATATAAGTCAATGAAGAGCCACTCTTCAAGAAAGTTCGC 120

QY 121 CGGTTTGTCTATCTTCCATCCAGTACCTGATATTTGGAAAATGTATAACAGGCACAG 180
Db 121 CGGTTTGTCTATCTTCCATCCAGTACCTGATATTTGGAAAATGTATAACAGGCACAG 180

QY 181 GCTTCTCTTGGACAGCAGAAAGAGTGCAGTCTTCAAGAGTCTCCCTCACTGGAACAAG 240
Db 181 GCTTCTCTTGGACAGCAGAAAGAGTGCAGTCTTCAAGAGTCTCCCTCACTGGAACAAG 240

QY 241 CTTTAAAGCAGATCAGAAAGTACTTCACTCTCTCACATCTTGGCAGCCAGTGAT 300
Db 241 CTTTAAAGCAGATCAGAAAGTACTTCACTCTCTCACATCTTGGCAGCCAGTGAT 300

QY 301 GGAATTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTCAGAGGCT 360
Db 301 GGAATTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGAGGTCAGAGGCT 360

QY 361 CGCTGTTTCTATCGCTTCAAAATTTCTCATCGAGATCTTCACTCAGAGATGACAGTTTG 420
Db 361 CGCTGTTTCTATCGCTTCAAAATTTCTCATCGAGATCTTCACTCAGAGATGACAGTTTG 420

QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTTGAA 480

QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGCCCTTGGATGATAGCAGATAGAAAA 540
Db 481 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGCCCTTGGATGATAGCAGATAGAAAA 540

QY 541 TCTTACTTTTGGGAAAAGAGTGGTGGCTTTGCTGTAGAGAGTTCCTTCTCAGGA 600
Db 541 TCTTACTTTTGGGAAAAGAGTGGTGGCTTTGCTGTAGAGAGTTCCTTCTCAGGA 600

QY 601 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGCTTATGCGAGGACTCACATTTTCC 660
Db 601 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGGCTTATGCGAGGACTCACATTTTCC 660

QY 661 AATGAACATCATCAGCAGAGATGAAGACCTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db |||||
QY 661 AATGAACATCATCAGCAGAGATGAAGACCTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db |||||
QY 721 TACTTAGTAATAAGCTTTCAGAGAAAGGCTCAGGAGATCAATTTGATGCTGTCAAA 780
Db |||||
QY 721 TACTTAGTAATAAGCTTTCAGAGAAAGGCTCAGGAGATCAATTTGATGCTGTCAAA 780
Db |||||
QY 781 ATTGAGCAGAGTTTAAACAGAGCCTTGCAGAGCTTGCAGTTGGCTCATTTGGAATGCAAT 840
Db |||||
QY 781 ATTGAGCAGAGTTTAAACAGAGCCTTGCAGAGCTTGCAGTTGGCTCATTTGGAATGCAAT 840
Db |||||
QY 841 TTGATGAACACAGTACATTCAGTTTGTAGCTGACAGATTAATTTGTGAACTTTGGAATTCCTCA 900
Db |||||
QY 841 TTGATGAACACAGTACATTCAGTTTGTAGCTGACAGATTAATTTGTGAACTTTGGAATTCCTCA 900
Db |||||
QY 901 AAGGTTTTTCAGGAGAAATCCTTTTGTATATGAAACATTTCTTTTGAAGGAA 960
Db |||||
QY 901 AAGGTTTTTCAGGAGAAATCCTTTTGTATATGAAACATTTCTTTTGAAGGAA 960
Db |||||
QY 961 ACAAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTTATGCGAGAAACC 1020
Db |||||
QY 961 ACAAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTTATGCGAGAAACC 1020
Db |||||
QY 1021 ACAGATAACGCTCTTCCACCTTTGGATGCGAGATTTT 1053
Db |||||
QY 1021 ACAGATAACGCTCTTCCACCTTTGGATGCGAGATTTT 1053
Db |||||

RESULT 2

US-09-787-491B-28
; Sequence 28, Application US/09787491B
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; TANG, Y. Tom;
; APPLICANT: CORLEY, Neil C.; GUEGLER, Karl J.;
; APPLICANT: GORONE, Gina A.; PATTERSON, Chandra;
; APPLICANT: HILLMAN, Jennifer L.; BAUGHN, Mariah R.;
; APPLICANT: LAL, Preeti; AZIMZAI, Yalda;
; APPLICANT: YOE, Henry; YANG, Junming
; TITLE OF INVENTION: RNA-ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0600 USN
; CURRENT APPLICATION NUMBER: US/09/787,491B
; PRIOR FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: PCT/US99/21688
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: US 60/128,660
; PRIOR FILING DATE: 1999-04-08
; PRIOR APPLICATION NUMBER: US 60/069,391
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: US 60/183,025
; PRIOR FILING DATE: 1998-09-22
; PRIOR APPLICATION NUMBER: US 60/155,246
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: US 09/158,720
; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PERL Program
; SEQ ID NO 28
; LENGTH: 2596
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 2604449CB1
US-09-787-491B-28

Query Match 100.0%; Score 1053; DB 34; Length 2596;
Best Local Similarity 100.0%; Pred. No. 2.9e-276;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGACCGCGAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
|||||

Db 56 ATGGCGACCGCGAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 115
QY 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCCTTAAGAAAGAGTTCTCGC 120
Db |||||
QY 116 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCCTTAAGAAAGAGTTCTCGC 175
Db |||||
QY 121 CGGTTTGTCTATCTTTTCCAAATCCAGTACCTTGATATTTGGAAAAATGTAATAACAGGCACAG 180
Db |||||
QY 176 CGGTTTGTCTATCTTTCCNAATCCAGTACCTTGATATTTGAAAAATGTAATAACAGGCACAG 235
Db |||||
QY 181 GCTTCCTTTGGAAGAGAGGTCGATTAATCAAGAGATCTCCCTCACTGGAACAG 240
Db |||||
QY 236 GCTTCCTTTGGAAGAGAGGTCGATTAATCAAGAGATCTCCCTCACTGGAACAG 295
Db |||||
QY 241 CTTAAAGCAGATCAGAGTACTTCTCTCACTCTAGCCTTTTGGACCCAGTCAAT 300
Db |||||
QY 296 CTTAAAGCAGATCAGAGTACTTCTCTCACTCTAGCCTTTTGGACCCAGTCAAT 355
Db |||||
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGAGGTCAGGTTCCAGAGGCT 360
Db |||||
QY 356 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGAGGTCAGGTTCCAGAGGCT 415
Db |||||
QY 361 CGCTGTTTCTATGCGCTTTCAAATTTCTATCAGAGATTTCTCACTCAGAGATGACAGTTTG 420
Db |||||
QY 416 CGCTGTTTCTATGCGCTTTCAAATTTCTATCAGAGATTTCTCACTCAGAGATGACAGTTTG 475
Db |||||
QY 421 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTCGAA 480
Db |||||
QY 476 CTGATAGACATTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTCGAA 535
Db |||||
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCTTTGCGATGATGAGAGATAGAAAA 540
Db |||||
QY 536 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGGCTTTGCGATGATGAGAGATAGAAAA 595
Db |||||
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTAGAGAGAGTCTTCTCTCAGGA 600
Db |||||
QY 596 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTAGAGAGAGTCTTCTCTCAGGA 655
Db |||||
QY 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGCTTATGCGAGAGTCACTTTTTC 660
Db |||||
QY 656 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGCTTATGCGAGAGTCACTTTTTC 715
Db |||||
QY 661 AATGAACATCATCAGCAGAGATGAAGACCTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db |||||
QY 716 AATGAACATCATCAGCAGAGATGAAGACCTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 775
Db |||||
QY 721 TACTTAGTAATAAGCTTTCAGAGAAAGGCTCAGGAGATCAATTTGATGCTGTCAAA 780
Db |||||
QY 776 TACTTAGTAATAAGCTTTCAGAGAAAGGCTCAGGAGATCAATTTGATGCTGTCAAA 835
Db |||||
QY 781 ATTGAGCAGAGTTTAAACAGAGCCTTGCAGTTGGCTCCTCAATTTGGAATGCAAT 840
Db |||||
QY 836 ATTGAGCAGAGTTTAAACAGAGCCTTGCAGTTGGCTCCTCAATTTGGAATGCAAT 895
Db |||||
QY 841 TTGATGAACACAGTACATTCAGTTTGTAGCTGACAGATTAATTTGGAAGTGGATTTCTCA 900
Db |||||
QY 896 TTGATGAACACAGTACATTCAGTTTGTAGCTGACAGATTAATTTGGAAGTGGATTTCTCA 955
Db |||||
QY 901 AAGGTTTTTCAGGAGAAATCCTTTTGTATTTTATGAAAAACATTTCTTTTGAAGGAAAA 960
Db |||||
QY 956 AAGGTTTTTCAGGAGAAATCCTTTTGTATTTTATGAAAAACATTTCTTTTGAAGGAAAA 1015
Db |||||
QY 961 ACAAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTTATGCGAGAAACC 1020
Db |||||
QY 1016 ACAAAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGTTTATGCGAGAAACC 1075
Db |||||
QY 1021 ACAGATAACGCTCTTCCACCTTTGGATGCGAGATTTT 1053
Db |||||
QY 1076 ACAGATAACGCTCTTCCACCTTTGGATGCGAGATTTT 1108
Db |||||

RESULT 3
US-60-128-660-20

; Sequence 20, Application US/60128660

; GENERAL INFORMATION:

; APPLICANT: Lal, Preeti

; APPLICANT: Tang, Y. Tom

; APPLICANT: Hillman, Jennifer L.

; APPLICANT: Yue, Henry

; APPLICANT: Yang, Junming

; APPLICANT: Baughn, Mariah R.

; TITLE OF INVENTION: RNA-ASSOCIATED PROTEINS

; FILE REFERENCE: PF-0682 P

; CURRENT APPLICATION NUMBER: US/60/128,660

; CURRENT FILING DATE: 1999-04-08

; NUMBER OF SEQ ID NOS: 26

; SOFTWARE: PERL Program

; SEQ ID NO 20

; LENGTH: 2596

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE: -

; OTHER INFORMATION: 2604449

; US-60-128-660-20

Query Match 100.0%; Score 1053; DB 81; Length 2596;

Best Local Similarity 100.0%; Pred. No. 2.9e-276;

Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGGGCGACCCGGAAGCGCGGAGCGGCTGGATCAGATCAGATCATCTTCA 60
DB 56 ATGGGCGACCCGGAAGCGCGGAGCGGCTGGATCAGATCAGATCATCTTCA 115
QY 61 GACACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
DB 116 GACACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 175
QY 121 CGGTTTGTCACTTTCCTCAATCCAGTACCTGATATTTGGAAATGATTAACACAGGCACAG 180
DB 176 CGGTTTGTCACTTTCCTCAATCCAGTACCTGATATTTGGAAATGATTAACACAGGCACAG 235
QY 181 GCTTCCTTCTGCACAGACAGAGGTCGACTTATCAAGAGATCTCCTCACTGGACACAG 240
DB 236 GCTTCCTTCTGCACAGACAGAGGTCGACTTATCAAGAGATCTCCTCACTGGACACAG 295
QY 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTCTAGCCCTTTTTCGAGCCAGTGAT 300
DB 296 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTCTAGCCCTTTTTCGAGCCAGTGAT 355
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
DB 356 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 415
QY 361 CGCTGTTTCTATGGCTTTTCAAAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
DB 416 CGCTGTTTCTATGGCTTTTCAAAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 475
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTTGAA 480
DB 476 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGAAATTTTATTAATGCAATTTGAA 535
QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGGCTTTCGATGGATAGCAGATAGAAA 540
DB 536 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGGCTTTCGATGGATAGCAGATAGAAA 595
QY 541 TCTACTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAGAGGATTTCTTCTCAGGA 600
DB 596 TCTACTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAGAGGATTTCTTCTCAGGA 655
QY 601 TCTTTTGTGCTATATTTCTGGCTTAAGAGAGAGGCTTATGCGAGGATCTCACTTTTCC 660
DB 656 TCTTTTGTGCTATATTTCTGGCTTAAGAGAGAGGCTTATGCGAGGATCTCACTTTTCC 715
QY 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATTTCCAA 720
DB 716 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATTTCCAA 775
```

```
QY 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGCTCAGGAGATCATTTGTTGATGCTGTCAAA 780
DB 776 TACTTAGTAAATAAGCCTTCAGAAAGAGGCTCAGGAGATCATTTGTTGATGCTGTCAAA 835
QY 781 ATTGAGCAGGAGTTTTTAAACAGAAAGCCTTGCCAGTTGGCCTCATTTGGAATGAATTCATT 840
DB 836 ATTGAGCAGGAGTTTTTAAACAGAAAGCCTTGCCAGTTGGCCTCATTTGGAATGAATTCATT 895
QY 841 TTGATGAAAACAGTACATTTGAGTTTGTAGCTGACAGATTACTTTGGAACCTTGGATTCTCA 900
DB 896 TTGATGAAAACAGTACATTTGAGTTTGTAGCTGACAGATTACTTTGGAACCTTGGATTCTCA 955
QY 901 AAGGTTTTTTCAGGCAGAAAATCCTTTTGAATTTTATGAAAACATTTCTTTAGAAAGGAAAA 960
DB 956 AAGGTTTTTTCAGGCAGAAAATCCTTTTGAATTTTATGAAAACATTTCTTTAGAAAGGAAAA 1015
QY 961 ACAAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTTCAGTTATGGCAGAAACC 1020
DB 1016 ACAAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTTCAGTTATGGCAGAAACC 1075
QY 1021 ACAGATAACGTCCTTCACTTGGATGAGATTTT 1053
DB 1076 ACAGATAACGTCCTTCACTTGGATGAGATTTT 1108
```

RESULT 4

US-10-170-235-20856

; Sequence 20856, Application US/10170235

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig

; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN

; TITLE OF INVENTION: TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF

; FILE REFERENCE: C1001380

; CURRENT APPLICATION NUMBER: US/10/170,235

; CURRENT FILING DATE: 2003-03-17

; NUMBER OF SEQ ID NOS: 42514

; SEQ ID NO 20856

; LENGTH: 3397

; TYPE: DNA

; ORGANISM: HUMAN

; US-10-170-235-20856

Query Match 100.0%; Score 1053; DB 49; Length 3397;

Best Local Similarity 100.0%; Pred. No. 3.2e-276;

Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGGGCGACCCGGAAGCGCGGAGCGGCTGGATCAGATCAGATCATCTTCA 60
DB 501 ATGGGCGACCCGGAAGCGCGGAGCGGCTGGATCAGATCAGATCATCTTCA 560
QY 61 GACACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
DB 561 GACACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 620
QY 121 CGGTTTGTCACTTTCCTCAATCCAGTACCTGATATTTGGAAATGATTAACAGGCACAG 180
DB 621 CGGTTTGTCACTTTCCTCAATCCAGTACCTGATATTTGGAAATGATTAACAGGCACAG 680
QY 181 GCTTCCTTCTGCACAGACAGAGGTCGACTTATCAAGAGATCTCCTCACTGGAACAG 240
DB 681 GCTTCCTTCTGCACAGACAGAGGTCGACTTATCAAGAGATCTCCTCACTGGAACAG 740
QY 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTCTAGCCCTTTTTCGAGCCAGTGAT 300
DB 741 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTCTAGCCCTTTTTCGAGCCAGTGAT 800
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
DB 801 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 860
QY 361 CGCTGTTTCTATGGCTTTTCAAAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
```

Db 861 CGCTGTTCTATGGCTTTCAAAATCTCATCGAAGATGTTCACTCAGAGATGATACAGTTTG 920
QY 421 CTGATAGACATTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAAATGCAATTTGAA 480
Db 921 CTGATAGACATTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTTAAATGCAATTTGAA 980
QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGGCTTCGATGGATAGCAGATAGAAAA 540
Db 981 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGGCTTCGATGGATAGCAGATAGAAAA 1040
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTGAAGAGAGTTTCTTCTCAGGA 600
Db 1041 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTGAAGAGAGTTTCTTCTCAGGA 1100
QY 601 TCTTTTGTCTGTATATCTCTGGCTTAAAGAAAGAGAGTCTTTATGCCAGGACTCACTTTTCC 660
Db 1101 TCTTTTGTCTGTATATCTCTGGCTTAAAGAAAGAGAGTCTTTATGCCAGGACTCACTTTTCC 1160
QY 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTCTTGGCTGTATGTTCCAA 720
Db 1161 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTCTTGGCTGTATGTTCCAA 1220
QY 721 TACTTAGTAATAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGATGCTGTCAAA 780
Db 1221 TACTTAGTAATAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGATGCTGTCAAA 1280
QY 781 ATTGAGCAGGAGTTTAAACAGAACCTTGCAGTGGCCTCATTTGGAATGCAATTTGCAAT 840
Db 1281 ATTGAGCAGGAGTTTAAACAGAACCTTGCAGTGGCCTCATTTGGAATGCAATTTGCAAT 1340
QY 841 TTGATGAACACAGTACATTTGAGTTTGTAGCTGACAGATTAATTTGTGGAACCTTGATTTCTCA 900
Db 1341 TTGATGAACACAGTACATTTGAGTTTGTAGCTGACAGATTAATTTGTGGAACCTTGATTTCTCA 1400
QY 901 AAGTTTTTTCAGGCGAAGAAATCCTTTTGAATTTATGGAACAATTTCTTTAGAAGGAAAA 960
Db 1401 AAGTTTTTTCAGGCGAAGAAATCCTTTTGAATTTATGGAACAATTTCTTTAGAAGGAAAA 1460
QY 961 ACNAATTTCTTGAGAAACGAGTTTCAGAGTATCAGCGTTTGGCAGTTTATGCGAAGAAC 1020
Db 1461 ACNAATTTCTTGAGAAACGAGTTTCAGAGTATCAGCGTTTGGCAGTTTATGCGAAGAAC 1520
QY 1021 ACAGATAACGTCCTTCACTTGGATCGAGATTTT 1053
Db 1521 ACAGATAACGTCCTTCACTTGGATCGAGATTTT 1553

RESULT 5

PCT-US02-18947-71
; Sequence 71, Application PC/TUS0218947
; GENERAL INFORMATION:
; APPLICANT: Rosetta Inpharmatics
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-228
; CURRENT APPLICATION NUMBER: PCT/US02/18947
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380, 770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 71
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AB036063
; DATABASE ENTRY DATE: 2001-06-18
PCT-US02-18947-71

Query Match 100.0%; Score 1053; DB 1; Length 4955;
Best Local Similarity 100.0%; Pred. No. 3.8e-276;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGGACCCGGAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 60

RESULT 6

Db 245 ATGGCGGACCCGGAAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
QY 61 GACACCAACGAAAGTGAATTAAGTCAATTAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 305 GACACCAACGAAAGTGAATTAAGTCAATTAAGAGCCACTCTTAAGAAAGAGTTCTCGC 364
QY 121 CGGTTTGTCTATCTTTCCAAATCCAGTACCTCTGATATTGGAAAAATGTATAAACAGGCACAG 180
Db 365 CGGTTTGTCTATCTTTCCAAATCCAGTACCTCTGATATTGGAAAAATGTATAAACAGGCACAG 424
QY 181 GCTTCCTCTCTGGAACAGCAGAGAGGTCGACTTATCAAGAGATCTCCCTCACTGCGAACAAAG 240
Db 425 GCTTCCTCTCTGGAACAGCAGAGAGGTCGACTTATCAAGAGATCTCCCTCACTGCGAACAAAG 484
QY 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTCTAGCCCTTTTTCAGCCAGTGTAT 300
Db 485 CTTAAAGCAGATGAGAAGTACTTCACTCTCACTCTAGCCCTTTTTCAGCCAGTGTAT 544
QY 301 GGAATTTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGTTCCAGAGGCT 360
Db 545 GGAATTTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGTTCCAGAGGCT 604
QY 361 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAAGTGTTCACCTCAGAGATGTACAGTTTG 420
Db 605 CGCTGTTTCTATGGCTTTCAAAATTTCTCATCGAAGTGTTCACCTCAGAGATGTACAGTTTG 664
QY 421 CTGATAGACATTTATCATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTTGAA 480
Db 665 CTGATAGACATTTATCATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATTTGAA 724
QY 481 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGGCTTCGATGGATAGCAGATAGAAAA 540
Db 725 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGGCTTCGATGGATAGCAGATAGAAAA 784
QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGTAGAAAGAGTGTCTTCTCTCAGGA 600
Db 785 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGTAGAAAGAGTGTCTTCTCTCAGGA 844
QY 601 TCTTTTGTCTATATTTCTGGCTTAAAGAAAGAGAGTCTTATGCCAGGACTCACATTTTTC 660
Db 845 TCTTTTGTCTATATTTCTGGCTTAAAGAAAGAGAGTCTTATGCCAGGACTCACATTTTTC 904
QY 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTGCTGATGTTCCAA 720
Db 905 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTGCTGATGTTCCAA 964
QY 721 TACTTAGTAATAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGTGATGCTGTCAAA 780
Db 965 TACTTAGTAATAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGTGATGCTGTCAAA 1024
QY 781 ATTGAGCAGGAGTTTAAACAGAAAGCCTTTGCCAGTTGGCTCATTTGGAATGAATTCAT 840
Db 1025 ATTGAGCAGGAGTTTAAACAGAAAGCCTTTGCCAGTTGGCTCATTTGGAATGAATTCAT 1084
QY 841 TTGATGAACACAGTACATTTGAGTTTGTAGCTGACAGATTAATTTGCGAACCTTGTGATTTCA 900
Db 1085 TTGATGAACACAGTACATTTGAGTTTGTAGCTGACAGATTAATTTGCGAACCTTGTGATTTCA 1144
QY 901 AAGGTTTTTTCAGGCGAAGAAATCCTTTTGAATTTTATGGAACAATTTCTTTAGAGGAAAA 960
Db 1145 AAGGTTTTTTCAGGCGAAGAAATCCTTTTGAATTTTATGGAACAATTTCTTTAGAGGAAAA 1204
QY 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTATTGCGCAGAAACC 1020
Db 1205 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGTATTGCGCAGAAACC 1264
QY 1021 ACAGATAACGTCCTTCACTTGGATGAGATTTT 1053
Db 1265 ACAGATAACGTCCTTCACTTGGATGAGATTTT 1297

```
PCT-US04-01100-71
; Sequence 71, Application PC/TUS0401100
; GENERAL INFORMATION:
; APPLICANT: Rosetta Inpharmatics LLC
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-228
; CURRENT APPLICATION NUMBER: PCT/US04/01100
; CURRENT FILING DATE: 2004-01-15
; PRIOR APPLICATION NUMBER: 10/342,887
; PRIOR FILING DATE: 2003-01-15
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 71
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US04-01100-71

Query Match      100.0%; Score 1053; DB 2; Length 4955;
Best Local Similarity 100.0%; Pred. No. 3.8e-276;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGGACCCGGGAAGCCGGGAAGCCGGCGGGCTGGATCAGATGAGAGATCATCTTCA 60
DB 245 ATGGCGGACCCGGGAAGCCGGGAAGCCGGCGGGCTGGATCAGATGAGAGATCATCTTCA 304
QY 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
DB 305 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 364
QY 121 CGGTTTGTCTATCTTTCCCAATCCAGTACCTGATATTTGGAAAATGTATATAACAGGACAG 180
DB 365 CGGTTTGTCTATCTTTCCCAATCCAGTACCTGATATTTGGAAAATGTATATAACAGGACAG 424
QY 181 GCTTCTCTCTGACAGACAGAGAGGTCGACTTATCAAAGGATCTCCCTCACTGGAACAAG 240
DB 425 GCTTCTCTCTGACAGACAGAGAGGTCGACTTATCAAAGGATCTCCCTCACTGGAACAAG 484
QY 241 CTATAAGCAGATGAGAAAGTACTTCTCATCTCTCATCTCTAGCCCTTTTTCGAGCCAGTGAT 300
DB 485 CTATAAGCAGATGAGAAAGTACTTCTCATCTCTCATCTCTAGCCCTTTTTCGAGCCAGTGAT 544

RESULT 7
US-10-172-118-71
; Sequence 71, Application US/10172118
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 71
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AB036063
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-71

Query Match      100.0%; Score 1053; DB 49; Length 4955;
Best Local Similarity 100.0%; Pred. No. 3.8e-276;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGGACCCGGGAAGCCGGGAAGCCGGCGGGCTGGATCAGATGAGAGATCATCTTCA 60
DB 245 ATGGCGGACCCGGGAAGCCGGGAAGCCGGCGGGCTGGATCAGATGAGAGATCATCTTCA 304
QY 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
DB 305 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 364
QY 121 CGGTTTGTCTATCTTTCCCAATCCAGTACCTGATATTTGGAAAATGTATATAACAGGACAG 180
DB 365 CGGTTTGTCTATCTTTCCCAATCCAGTACCTGATATTTGGAAAATGTATATAACAGGACAG 424
QY 181 GCTTCTCTCTGACAGACAGAGAGGTCGACTTATCAAAGGATCTCCCTCACTGGAACAAG 240
DB 425 GCTTCTCTCTGACAGACAGAGAGGTCGACTTATCAAAGGATCTCCCTCACTGGAACAAG 484
QY 241 CTATAAGCAGATGAGAAAGTACTTCTCATCTCTCATCTCTAGCCCTTTTTCGAGCCAGTGAT 300
DB 485 CTATAAGCAGATGAGAAAGTACTTCTCATCTCTCATCTCTAGCCCTTTTTCGAGCCAGTGAT 544
```

QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTCAGAGGTCAGAGGCT 360
Db |||||
QY 545 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTCAGAGGTCAGAGGCT 604
Db |||||
QY 361 CGCTGTTCTATGGCTTTTCAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db |||||
QY 605 CGCTGTTCTATGGCTTTTCAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 664
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 480
Db |||||
QY 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 724
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGATGGATAGCAGATAGAAA 540
Db |||||
QY 725 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGATGGATAGCAGATAGAAA 784
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAGAGATTTTCTCTCAGGA 600
Db |||||
QY 785 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAGAGATTTTCTCTCAGGA 844
QY 601 TCTTTTGTCTGTATATTTCTGGCTTAAGAAAGAGAGGTCCTTATGCCAGACTCTCTTTCC 860
Db |||||
QY 845 TCTTTTGTCTGTATATTTCTGGCTTAAGAAAGAGAGGTCCTTATGCCAGACTCTCTTTCC 904
QY 661 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTTGTCTTGCCTGATGTTCCAA 720
Db |||||
QY 905 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTTGTCTTGCCTGATGTTCCAA 964
QY 721 TACTTAGTAAATPAGCTTTTCAAGAAAGGGTCAGGAGATCAATTTGTATGCTGTCAAA 780
Db |||||
QY 965 TACTTAGTAAATPAGCTTTTCAAGAAAGGGTCAGGAGATCAATTTGTATGCTGTCAAA 1024
QY 781 ATTGAGCAGGAGTTTAAACAGAGCTTGCAGATTTGGCTCATTTGGAATGAATTCATT 840
Db |||||
QY 1025 ATTGAGCAGGAGTTTAAACAGAGCTTGCAGATTTGGCTCATTTGGAATGAATTCATT 1084
QY 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATTTGTGAACTTGTGATCTCA 900
Db |||||
QY 1085 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATTTGTGAACTTGTGATCTCA 1144
QY 901 AAGGTTTTTTCAGGAGAAATCCTTTTGAATTTATGGAAGAAATTTCTTTAGAGGAAAA 960
Db |||||
QY 1145 AAGGTTTTTTCAGGAGAAATCCTTTTGAATTTATGGAAGAAATTTCTTTAGAGGAAAA 1204
QY 961 ACAATTTCTTTGGAAGACAGTTTTCAGAGTATCAGCGTTTTCAGATTTATGCGAGAAACC 1020
Db |||||
QY 1205 ACAATTTCTTTGGAAGACAGTTTTCAGAGTATCAGCGTTTTCAGATTTATGCGAGAAACC 1264
QY 1021 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1053
Db |||||
QY 1265 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1297

RESULT 8
US-10-342-887-71
; Sequence 71, Application US/10342887
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14

; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 71
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-71

Query Match 100.0%; Score 1053; DB 52; Length 4955;
Best Local Similarity 100.0%; Pred. No. 3.8e-276;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAGCGCGAGCGCGGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db |||||
QY 245 ATGGGCGACCCGGAAGCGCGAGCGCGGGCTGGATCAGGATGAGAGATCATCTTCA 304
Db |||||
QY 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGGCACCTCCTAAGAAAGAGTTCTCGC 120
Db |||||
QY 305 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGGCACCTCCTAAGAAAGAGTTCTCGC 364
QY 121 CGGTTTGTCTATCTTTCCAATCCAGTACCTCTGATATTTGGAAAAATGTATAAACAGGCACAG 180
Db |||||
QY 365 CGGTTTGTCTATCTTTCCAATCCAGTACCTCTGATATTTGGAAAAATGTATAAACAGGCACAG 424
QY 181 GCTTCCTTCTGGACAGCAGAGAGGTCGACTTCAATCAAGAGATCTCCCTCCTCAGTGGACAA 240
Db |||||
QY 425 GCTTCCTTCTGGACAGCAGAGAGGTCGACTTCAATCAAGAGATCTCCCTCCTCAGTGGACAA 484
QY 241 CTTTAAAGCAGATCAGAGATGACTTTCATCTCTCACATCTTAGCCTTTTGTGACAGCAGATGAT 300
Db |||||
QY 485 CTTTAAAGCAGATCAGAGATGACTTTCATCTCTCACATCTTAGCCTTTTGTGACAGCAGATGAT 544
QY 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTCAGAGGTCAGAGGCT 360
Db |||||
QY 545 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTCAGAGGTCAGAGGCT 604
QY 361 CGCTGTTTCTATGGCTTTTCAATTCATCAGAGATGTTTCACTCAGAGATGTACAGTTTG 420
Db |||||
QY 605 CGCTGTTTCTATGGCTTTTCAATTCATCAGAGATGTTTCACTCAGAGATGTACAGTTTG 664
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 480
Db |||||
QY 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 724
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGATGGATAGCAGATAGAAA 540
Db |||||
QY 725 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTCGATGGATAGCAGATAGAAA 784
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAGAGATTTTCTCTCAGGA 600
Db |||||
QY 785 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAGAGATTTTCTCTCAGGA 844
QY 601 TCTTTTGTCTGTATATTTCTGGCTTAAGAAAGAGAGGTCCTTATGCCAGACTCTCTTTCC 860
Db |||||
QY 845 TCTTTTGTCTGTATATTTCTGGCTTAAGAAAGAGAGGTCCTTATGCCAGACTCTCTTTCC 904
QY 661 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTTGTCTTGCCTGATGTTCCAA 720
Db |||||
QY 905 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTTGTCTTGCCTGATGTTCCAA 964
QY 721 TACTTAGTAAATPAGCTTTTCAAGAAAGGGTCAGGAGATCAATTTGTATGCTGTCAAA 780
Db |||||
QY 965 TACTTAGTAAATPAGCTTTTCAAGAAAGGGTCAGGAGATCAATTTGTATGCTGTCAAA 1024
QY 781 ATTGAGCAGGAGTTTAAACAGAGCTTGCAGATTTGGCTCATTTGGAATGAATTCATT 840
Db |||||
QY 1025 ATTGAGCAGGAGTTTAAACAGAGCTTGCAGATTTGGCTCATTTGGAATGAATTCATT 1084
QY 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATTTGTGAACTTGTGATCTCA 900
Db |||||
QY 1085 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATTTGTGAACTTGTGATCTCA 1144

QY 901 AAGGTTTTTCAGGCGAGAAAATCCTTTTGATTTTATGGAAGAAACATTTCTTTAGAGGAAAA 960
DB 1145 AAGGTTTTTCAGGCGAGAAAATCCTTTTGATTTTATGGAAGAAACATTTCTTTAGAGGAAAA 1204
QY 961 ACAAATTTCTTGAGAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAACCC 1020
DB 1205 ACAAATTTCTTGAGAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGCGAGAACCC 1264
QY 1021 ACAGATAACGCTTTCACCTTTGATGCGAGATTTT 1053
DB 1265 ACAGATAACGCTTTCACCTTTGATGCGAGATTTT 1297

RESULT 9

US-10-698-228-3
; Sequence 3, Application US/10698228
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619WOOP
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 3
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-3

Query Match 100.0%; Score 1053; DB 58; Length 4955;
Best Local Similarity 100.0%; Pred. No. 3.8e-276;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGGAAGCGCGGAAGCGCGGCGTGGATCAGATCAGAGATCATCTTCA 60
DB 245 ATGGGCGACCCGGAAGCGCGGAAGCGCGGCGTGGATCAGATCAGAGATCATCTTCA 304
QY 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
DB 305 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 364
QY 121 CGGTTTGTCATCTTCCCAATCCAGTACCTGATATTTGGAATGTATAAAGAGGACAG 180
DB 365 CGGTTTGTCATCTTCCCAATCCAGTACCTGATATTTGGAATGTATAAAGAGGACAG 424
QY 181 GCTTCCTCTTGACAGCAGAGAGGTCGACTTATCAAGAGTCTCCCTCACTGGAACNAG 240
DB 425 GCTTCCTCTTGACAGCAGAGAGGTCGACTTATCAAGAGTCTCCCTCACTGGAACNAG 484
QY 241 CTTAAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTAGCCCTTTTTCGAGCCAGTGAT 300
DB 485 CTTAAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTAGCCCTTTTTCGAGCCAGTGAT 544
QY 301 GGAATTTGAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
DB 545 GGAATTTGAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 604
QY 361 CGCTGTTCTATGGCTTTCAAAATTCATCAGAGATGTTTCACTCAGAGATGTACAGTTTG 420
DB 605 CGCTGTTCTATGGCTTTCAAAATTCATCAGAGATGTTTCACTCAGAGATGTACAGTTTG 664
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480

DB 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 724
QY 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCCCTTGGATGATGATAGCAGATAGAAA 540
DB 725 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCCCTTGGATGATGATAGCAGATAGAAA 784
QY 541 TCTACTTTTGGGAAAAAGAGTGGGCTTTGCTGCTGTAAGAGGAGTTTCTTCTCAGGA 600
DB 785 TCTACTTTTGGGAAAAAGAGTGGGCTTTGCTGCTGTAAGAGGAGTTTCTTCTCAGGA 844
QY 601 TCTTTTGGCTGCTATATCTGGCTAAAGAAAGAGAGTCTTATGCCAGACTCACATTTTCC 660
DB 845 TCTTTTGGCTGCTATATCTGGCTAAAGAAAGAGAGTCTTATGCCAGACTCACATTTTCC 904
QY 661 AATGAACTCATCAGCAGAGATGAAGACTTCACTGTGACTTTTGGCTGCTGATGTTCAA 720
DB 905 AATGAACTCATCAGCAGAGATGAAGACTTCACTGTGACTTTTGGCTGCTGATGTTCAA 964
QY 721 TACTTAGTAAATAAGCCTTCAGAAAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAA 780
DB 965 TACTTAGTAAATAAGCCTTCAGAAAGAAAGGTCAGGAGATCATTTGTGATGCTGTCAA 1024
QY 781 ATTGACGAGAGTTTAAACAGAGCCTTCCAGTTCGCTCATTTGGAATGAATGCAATT 840
DB 1025 ATTGACGAGAGTTTAAACAGAGCCTTCCAGTTCGCTCATTTGGAATGAATGCAATT 1084
QY 841 TTGATGAAACAGTACATTTGAGTTTGTAGCTGACAGATTACTTGTGAACTTTGGATTCTCA 900
DB 1085 TTGATGAAACAGTACATTTGAGTTTGTAGCTGACAGATTACTTGTGAACTTTGGATTCTCA 1144
QY 901 AAGGTTTTTCAGGCGAGAAAATCCTTTTGATTTTATGAAAAACATTTCTTTAAGAGGAAAA 960
DB 1145 AAGGTTTTTCAGGCGAGAAAATCCTTTTGATTTTATGAAAAACATTTCTTTAAGAGGAAAA 1204
QY 961 ACAAATTTCTTGAGAACGAGTTTCAGAGTATCAGCGTTTTCAGTATGCGCAGAAACC 1020
DB 1205 ACAAATTTCTTGAGAACGAGTTTTCAGAGTATCAGCGTTTTCAGTATGCGCAGAAACC 1264
QY 1021 ACAGATAACGCTTTCACCTTTGGATGCGAGATTTT 1053
DB 1265 ACAGATAACGCTTTCACCTTTGGATGCGAGATTTT 1297

RESULT 10

US-10-990-328-49
; Sequence 49, Application US/10990328
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 1849
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-990-328-49

Query Match 100.0%; Score 1052.6; DB 67; Length 1849;
Best Local Similarity 99.9%; Pred. No. 3.2e-276;
Matches 1052; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGCGACCCGGAAGCGCGGAAGCGCGGCTGGATCAGAGTGAAGATCATCTTCA 60
DB 85 ATGGGCGACCCGGAAGCGCGGAAGCGCGGCTGGATCAGAGTGAAGATCATCTTCA 144
QY 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
DB 145 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 204

```
QY 121 CGGTTTGTCATCTTTCCAAATCCAGTACCTCGATATTTGGAAATGTATAAACAGGCACAG 180
D 205 CGGTTTGTCATCTTTCCAAATCCAGTACCTCGATATTTGGAAATGTATAAACAGGCACAG 264
QY 181 GCTTCTCTTGACAGCAGAGAGTGCAGTATTAAGAGGATCTCCCTCACTGGAAACAAG 240
D 265 GCTTCTCTTGACAGCAGAGAGTGCAGTATTAAGAGGATCTCCCTCACTGGAAACAAG 324
QY 241 CTTAAGCAGATGAGAGTACCTCATCTCATCTCTAGCCCTTTTTCGACGCCAGTGAT 300
D 325 CTTAAGCAGATGAGAGTACCTCATCTCATCTCTAGCCCTTTTTCGACGCCAGTGAT 384
QY 301 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
D 385 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 444
QY 361 CGCTGTTCTATGCGCTTTCAAAATCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 420
D 445 CGCTGTTCTATGCGCTTTCAAAATCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 504
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGATTTTATTAATGCAATTGAA 480
D 505 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGATTTTATTAATGCAATTGAA 564
QY 481 ACCATGCCCTATGTTAAGAAAGAAAGCAGATTCGGGCTTTCGATGGATAGCAGATAGAAA 540
D 565 ACCATGCCCTATGTTAAGAAAGAAAGCAGATTCGGGCTTTCGATGGATAGCAGATAGAAA 624
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTAGAGAGGATTTTCTTCAGGA 600
D 625 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTAGAGAGGATTTTCTTCAGGA 684
QY 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGCTTTATGCCAGGACTCATCTTTTCC 660
D 685 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGCTTTATGCCAGGACTCATCTTTTCC 744
QY 661 AATGAACACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTTGCCTGATGTTCAA 720
D 745 AATGAACACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTTGCCTGATGTTCAA 804
QY 721 TACTTAGTAAATTAAGCCTTTTCAAGAAAGGGGTCCAGGAGATCAATTTGTGATGTTCAA 780
D 805 TACTTAGTAAATTAAGCCTTTTCAAGAAAGGGGTCCAGGAGATCAATTTGTGATGTTCAA 864
QY 781 ATTGACGAGGAGTGTATTAAGCAGAGCCTTGCAGTTGGCTCATTTGGAATGGAATTGCAAT 840
D 865 ATTGACGAGGAGTGTATTAAGCAGAGCCTTGCAGTTGGCTCATTTGGAATGGAATTGCAAT 924
QY 841 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTAATCTTGCGAATCTGATCTCA 900
D 925 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTAATCTTGCGAATCTGATCTCA 984
QY 901 AAGGTTTTTTCAGGAGAGAAATCTTTTGAATTTATGGAAGAAACATTTCTTTAGAGGAGAAA 960
D 985 AAGGTTTTTTCAGGAGAGAAATCTTTTGAATTTATGGAAGAAACATTTCTTTAGAGGAGAAA 1044
QY 961 ACAAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGAGTTATGCGAGAAACC 1020
D 1045 ACAAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGAGTTATGCGAGAAACC 1104
QY 1021 ACAGATAACGCTTTCACCTTGGATGATGATTTT 1053
D 1105 ACAGATAACGCTTTCACCTTGGATGATGATTTT 1137
```

RESULT 11

```
US-60-505-218-20
; Sequence 20, Application US/60505218
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CANCER, METHODS OF DETECTION AND USES THEREOF
```

```
; FILE REFERENCE: CL001482
; CURRENT APPLICATION NUMBER: US/60/505,218
; CURRENT FILING DATE: 2003-09-24
; NUMBER OF SEQ ID NOS: 22507
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 1849
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-60-505-218-20
```

```
Query Match 100.0%; Score 1052.6; DB 120; Length 1849;
Best Local Similarity 99.9%; Pred. No. 3.2e-276;
Matches 1052; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ATGGGGCGACCCGGAAAGCGCGGAGCGGCGGCTGGATCAGATGAGAGATCATCTTCA 60
D 85 ATGGGGCGACCCGGAAAGCGCGGAGCGGCGGCTGGATCAGATGAGAGATCATCTTCA 144
QY 61 GACACCAACGAAAGTCAAAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
D 145 GACACCAACGAAAGTCAAAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 204
QY 121 CGGTTTGTCTATCTTTCCAAATCCAGTACCTCGATATTTGGAAATGTATTAACAGGCACAG 180
D 205 CGGTTTGTCTATCTTTCCAAATCCAGTACCTCGATATTTGGAAATGTATTAACAGGCACAG 264
QY 181 GCTTCTCTTGACAGCAGAGAGTGCAGTATTAAGAGATCTCCCTCACTGGAACAAG 240
D 265 GCTTCTCTTGACAGCAGAGAGTGCAGTATTAAGAGATCTCCCTCACTGGAACAAG 324
QY 241 CTTAAGCAGATGAGAGTACCTCATCTCTCATCTCTAGCCCTTTTTCGAGCCAGTGAT 300
D 325 CTTAAGCAGATGAGAGTACCTCATCTCTCATCTCTAGCCCTTTTTCGAGCCAGTGAT 384
QY 301 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
D 385 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 444
QY 361 CGCTGTTCTATGCGCTTTCAAAATCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 420
D 445 CGCTGTTCTATGCGCTTTCAAAATCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 504
QY 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGATTTTATTAATGCAATTGAA 480
D 505 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGATTTTATTAATGCAATTGAA 564
QY 481 ACCATGCCCTATGTTAAGAAAGAAAGCAGATTCGGGCTTTCGATGGATAGCAGATAGAAA 540
D 565 ACCATGCCCTATGTTAAGAAAGAAAGCAGATTCGGGCTTTCGATGGATAGCAGATAGAAA 624
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGTCTGTAGAGAGGATTTTCTTCAGGA 600
D 625 TCTACTTTTGGGAAAGAGTGGTGGCTTTTGTCTGTAGAGAGGATTTTCTTCAGGA 684
QY 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGCTTTATGCCAGGACTCATCTTTTCC 660
D 685 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGCTTTATGCCAGGACTCATCTTTTCC 744
QY 661 AATGAACACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTTGCCTGATGTTCAA 720
D 745 AATGAACACTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTTGTGCTTGCCTGATGTTCAA 804
QY 721 TACTTAGTAAATTAAGCCTTTTCAAGAAAGGGGTCCAGGAGATCAATTTGTGATGTTCAA 780
D 805 TACTTAGTAAATTAAGCCTTTTCAAGAAAGGGGTCCAGGAGATCAATTTGTGATGTTCAA 864
QY 781 ATTGACGAGGAGTGTATTAAGCAGAGCCTTGCAGTTGGCTCATTTGGAATGGAATTGCAAT 840
D 865 ATTGACGAGGAGTGTATTAAGCAGAGCCTTGCAGTTGGCTCATTTGGAATGGAATTGCAAT 924
QY 841 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTAATCTTGCGAATCTGATCTCA 900
```

Db 925 TTGATGAAACAGTACATTCAGTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 984
QY 901 AAGGTTTTTCAGCGCAGAAATCCTTTTGATTTTATGGAACAATTTCTTTAGAAAGAAA 960
Db 985 AAGGTTTTTCAGCGCAGAAATCCTTTTGATTTTATGGAACAATTTCTTTAGAAAGAAA 1044
QY 961 ACAAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTGTCAGTTATGCGAGAACC 1020
Db 1045 ACAAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTGTCAGTTATGCGAGAACC 1104
QY 1021 ACAGATAACGTCCTTCACTTGGATGAGATTTT 1053
Db 1105 ACAGATAACGTCCTTCACTTGGATGAGATTTT 1137

RESULT 12
US-60-449-629-12
; Sequence 12, Application US/60449629
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CANCER, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001449
; CURRENT APPLICATION NUMBER: US/60/449,629
; NUMBER OF SEQ ID NOS: 2712
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 3397
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-449-629-12

Query Match 100.0%; Score 1052.6; DB 114; Length 3397;
Best Local Similarity 99.9%; Pred. No. 4.2e-276;
Matches 1052; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGCGACCCGAAAGCGCGAAGCGCGCGCTGGATCAGATCAGATCATCTTCA 60
Db 501 ATGGGCGACCCGAAAGCGCGAAGCGCGCGCTGGATCAGATCAGATCATCTTCA 560
QY 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
Db 561 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 620
QY 121 CGGTTTGTCACTTTTCCAAATCCAGTACCTGTATATTTGGAAATGTATAAACAGGCACAG 180
Db 621 CGGTTTGTCACTTTTCCAAATCCAGTACCTGTATATTTGGAAATGTATAAACAGGCACAG 680
QY 181 GCTTCCTCTTGACACGACGAGGTCGACTTATCAAGGATCTCCCTCACTGGAACAAG 240
Db 681 GCTTCCTCTTGACACGACGAGGTCGACTTATCAAGGATCTCCCTCACTGGAACAAG 740
QY 241 CTTAAAGCAGATGAGAAAGTACTTCACTCTCACTCTAGACCTTTTTCAGCCAGTGAT 300
Db 741 CTTAAAGCAGATGAGAAAGTACTTCACTCTCACTCTAGACCTTTTTCAGCCAGTGAT 800
QY 301 GGAATTTGTAATGAAATTTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 360
Db 801 GGAATTTGTAATGAAATTTTGGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 860
QY 361 CGCTGTTCTATGGCTTTCAAAATCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 420
Db 861 CGCTGTTCTATGGCTTTCAAAATCTCATCGAGAAATGTTCACTCAGAGATGTACAGTTTG 920
QY 421 CTGATAGACACTTACATCAGAGATCCAAAGAAAGGAAATTTTATTTAATGCAATTTGAA 480
Db 921 CTGATAGACACTTACATCAGAGATCCAAAGAAAGGAAATTTTATTTAATGCAATTTGAA 980
QY 481 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGGCCCTTCGATGGATAGCAGATAGAAA 540
Db 981 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGGCCCTTCGATGGATAGCAGATAGAAA 1040

QY 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAAGAGATTTTCTTCTCAGGA 600
Db 1041 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTGTAGAAGAGATTTTCTTCTCAGGA 1100
QY 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAAAGAGAGGTCTTATGCCAGGACTCACATTTTCC 660
Db 1101 TCTTTTGTGCTATATTTCTGGCTTAAAGAAAGAGAGGTCTTATGCCAGGACTCACATTTTCC 1160
QY 661 AATGAACCTCATCAGCAGAGATGAAGACTTTCACCTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 1161 AATGAACCTCATCAGCAGAGATGAAGACTTTCACCTGTGACTTTGCTTGCCTGATGTTCCAA 1220
QY 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGCTCAGGAGAGTCAFTTGTGATGCTGTCAAA 780
Db 1221 TACTTAGTAAATAAGCCTTCAGAAAGAGGCTCAGGAGAGTCAFTTGTGATGCTGTCAAA 1280
QY 781 ATTGAGCAGGAGTTTTTAAACAGAGCCTTCCAGAGTTGGCCTCATTTGGAATGAATTCAT 840
Db 1281 ATTGAGCAGGAGTTTTTAAACAGAGCCTTCCAGAGTTGGCCTCATTTGGAATGAATTCAT 1340
QY 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
Db 1341 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 1400
QY 901 AAGGTTTTTTCAGGCAGAAAATCCTTTTGATTTTATGGAACAATTTCTTTAGAAAGAAA 960
Db 1401 AAGGTTTTTTCAGGCAGAAAATCCTTTTGATTTTATGGAACAATTTCTTTAGAAAGAAA 1460
QY 961 ACAAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTTGCAGTTATGGCAGAAACC 1020
Db 1461 ACAAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTTGCAGTTATGGCAGAAACC 1520
QY 1021 ACAGATAACGTCCTTCACTTGGATGAGATTTT 1053
Db 1521 ACAGATAACGTCCTTCACTTGGATGAGATTTT 1553

RESULT 13

US-10-698-228-12
; Sequence 12, Application US/10698228
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New protein and its DNA
; FILE REFERENCE: 2619W00P
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 12
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-12

Query Match 99.8%; Score 1051.4; DB 58; Length 1053;
Best Local Similarity 99.9%; Pred. No. 5.4e-276;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGGCGACCCGAAAGCGCGAAGCGCGCGCTGGATCAGATGAGATCATCTTCA 60
Db 1 ATGGGCGACCCGAAAGCGCGAAGCGCGCGCTGGATCAGATGAGATCATCTTCA 60
QY 61 GACACCAACGAAAGTGAATAAAGTCAAAATGAAGAGCCACTTCTTAAGAAAGAGTTCTCGC 120

Db 61 GACACCAAGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
QY 121 CGGTTTGTCTATCTTTCCAAATCCAGTACCCTGATATTTGGAAATGATATAACAGGCACAG 180
Db 121 CGGTTTGTCTATCTTTCCAAATCCAGTACCCTGATATTTGGAAATGATATAACAGGCACAG 180
QY 181 GCTTCCTCTGGACACAGAGAGGTCGACTTATCAAAAGGATCTCCCTCACTGGAACAAG 240
Db 181 GCTTCCTCTGGACACAGAGAGGTCGACTTATCAAAAGGATCTCCCTCACTGGAACAAG 240
QY 241 CTTAAGCAGATGAGAGTACTTCTATCTCTACATCTTAGCCCTTTTTCGAGCCAGTGAT 300
Db 241 CTTAAGCAGATGAGAGTACTTCTATCTCTACATCTTAGCCCTTTTTCGAGCCAGTGAT 300
QY 301 CGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGCT 360
Db 301 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGCT 360
QY 361 CGCTGTTCTATGGCTTTCAAAATTTCTCATCGAATGTTTCACTCAGAGATGTACAGTTTG 420
Db 361 CGCTGTTCTATGGCTTTCAAAATTTCTCATCGAATGTTTCACTCAGAGATGTACAGTTTG 420
QY 421 CTGATAGACACTTACATCAGAGATCCAAAGAAAGGGAATTTTATTAATGCAATTGAA 480
Db 421 CTGATAGACACTTACATCAGAGATCCAAAGAAAGGGAATTTTATTAATGCAATTGAA 480
QY 481 ACCATGCCCTATGTTAAGAAAGAGAGATTTGGCCCTTGGATGATAGATAGAGAA 540
Db 481 ACCATGCCCTATGTTAAGAAAGAGAGATTTGGCCCTTGGATGATAGATAGAGAA 540
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAAAGGAGTTTCTTCTCAGGA 600
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAAAGGAGTTTCTTCTCAGGA 600
QY 601 TCTTTTGTCTATATTTCTGGCTTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTCC 660
Db 601 TCTTTTGTCTATATTTCTGGCTTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTCC 660
QY 661 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTGCTGTGATGATGTTCCAA 720
Db 661 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTGCTGTGATGATGTTCCAA 720
QY 721 TACTTAGTAAATAAGCTTTCAGAGAAAGGTCAGGAGATCATTTGTCATCTGTCAA 780
Db 721 TACTTAGTAAATAAGCTTTCAGAGAAAGGTCAGGAGATCATTTGTCATCTGTCAA 780
QY 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGGCAGTTGGCTCATTTGGAATGAATTCAT 840
Db 781 ATTGAGCAGGAGTTTAAACAGAGCCTTGGCAGTTGGCTCATTTGGAATGAATTCAT 840
QY 841 TTGATGAACAGTACATTTAGTTTGTAGCTGACAGATTAATTTGGAATTCATTCAT 900
Db 841 TTGATGAACAGTACATTTAGTTTGTAGCTGACAGATTAATTTGGAATTCATTCAT 900
QY 901 AAGGTTTTTCAGGAGAAATCTTTTGAATTTTATGGAATTCATTTTGAAGGAAA 960
Db 901 AAGGTTTTTCAGGAGAAATCTTTTGAATTTTATGGAATTCATTTTGAAGGAAA 960
QY 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGAGTATGCGA 1020
Db 961 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGAGTATGCGA 1020
QY 1021 ACAGATAAGCTTTCACCTTGGATCAGATTTT 1053
Db 1021 ACAGATAAGCTTTCACCTTGGATCAGATTTT 1053

RESULT 14

US-10-698-228-4

; Sequence 4, Application US/10698228

; GENERAL INFORMATION:

; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA

; TITLE OF INVENTION: New Protein and its DNA

FILE REFERENCE: 2619WOOP
CURRENT APPLICATION NUMBER: US/10/698,228
CURRENT FILING DATE: 2003-10-30
PRIOR APPLICATION NUMBER: US/10/019,733
PRIOR FILING DATE: 2001-12-28
PRIOR APPLICATION NUMBER: JP 11-181131
PRIOR FILING DATE: 1999-06-28
PRIOR APPLICATION NUMBER: JP 11-192391
PRIOR FILING DATE: 1999-07-06
PRIOR APPLICATION NUMBER: JP 2000-017770
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 14
SEQ ID NO 4
LENGTH: 1081
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION:
US-10-698-228-4

Query Match 99.8%; Score 1051.4; DB 58; Length 1081;
Best Local Similarity 99.9%; Pred. No. 5.5e-276;
Matches 1052; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGGCGACCCGAAAGCCGCGGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 60
Db 20 ATGGGCGACCCGAAAGCCGCGGCGGCGGCTGGATCAGGATGAGAGATCATCTTCA 79
QY 61 GACACCAAGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 80 GACACCAAGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 139
QY 121 CGGTTTGTCTATCTTTCCAAATCCAGTACCCTGATATTTGGAAATGATATAACAGGCACAG 180
Db 140 CGGTTTGTCTATCTTTCCAAATCCAGTACCCTGATATTTGGAAATGATATAACAGGCACAG 199
QY 181 GCTTCCTCTGGACACAGAGAGTTCGACTTATCAAGGATCTCCCTCACTGGAACAAG 240
Db 200 GCTTCCTCTGGACACAGAGAGTTCGACTTATCAAGGATCTCCCTCACTGGAACAAG 259
QY 241 CTTAAGCAGATGAGAGTACTTCTCATCTCACTCTTAGCCCTTTTTCGAGCCAGTGAT 300
Db 260 CTTAAGCAGATGAGAGTACTTCTCATCTCTCACTCTTAGCCCTTTTTCGAGCCAGTGAT 319
QY 301 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGCT 360
Db 320 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGCT 379
QY 361 CGCTGTTCTATGGCTTTCAAAATTTCTCATCGAATGTTTCACTCAGAGATGTACAGTTG 420
Db 380 CGCTGTTCTATGGCTTTCAAAATTTCTCATCGAATGTTTCACTCAGAGATGTACAGTTG 439
QY 421 CTGATAGACACTTACATCAGAGATCCAAAGAAAGGGAATTTTATTAATGCAATTGAA 480
Db 440 CTGATAGACACTTACATCAGAGATCCAAAGAAAGGGAATTTTATTAATGCAATTGAA 499
QY 481 ACCATGCCCTATGTTAAGAAAGAGAGATTTGGCCCTTGGATGATAGATAGAGAA 540
Db 500 ACCATGCCCTATGTTAAGAAAGAGAGATTTGGCCCTTGGATGATAGAGAA 559
QY 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAGAGAGTTTCTTCTCAGGA 600
Db 560 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAGAGAGAGTTTCTTCTCAGGA 619
QY 601 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGTTTATGCCAGGACTCACTTTTTC 660
Db 620 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGTTTATGCCAGGACTCACTTTTTC 679
QY 661 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTTGTGCTGCTGATGTTCCAA 720
Db 680 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTTGTGCTGCTGATGTTCCAA 739
QY 721 TACTTAGTAAATAAGCCTTTCAGAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAA 780

Db 740 TACTTAGTAAATAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGTCATGCTGTCAAA 799
Qy 781 ATTGACGAGGTTTTTAACAGAGCCCTTGCAGTTGGCCTCATTTGGAATGAATTCATT 840
Db 800 ATTGACGAGGTTTTTAACAGAGCCCTTGCAGTTGGCCTCATTTGGAATGAATTCATT 859
Qy 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
Db 860 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 919
Qy 901 AAGGTTTTTCAGCGAGAAAATCCTTTTGTATTTATGAGAAAACATTTCTTTAGAGGAAAA 960
Db 920 AAGGTTTTTCAGCGAGAAAATCCTTTTGTATTTATGAGAAAACATTTCTTTAGAGGAAAA 979
Qy 961 ACAATATTTCTTTGAGAAAACAGTTTCAGAGTATCAGCGTTTTCAGAGTTATGGCAGAAACC 1020
Db 980 ACAATATTTCTTTGAGAAAACAGTTTCAGAGTATCAGCGTTTTCAGAGTTATGGCAGAAACC 1039
Qy 1021 ACAGATAACGCTTCTACACCTTGGATGCAGATTTT 1053
Db 1040 ACAGATAACGCTTCTACCTTGGATGCAGATTTT 1072

RESULT 15
US-09-629-469A-12810
; Sequence 12810, Application US/09629469A
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: SAITO, KAORU
; APPLICANT: YAMAMOTO, JUNICHI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: WAKAMATSU, AI
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: OTSUKI, TETSUJI
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE
; FILE REFERENCE: 084335/0123
; CURRENT APPLICATION NUMBER: US/09/629,469A
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: JP 1999-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 1999-300253
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/183,322
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 19025
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12810
; LENGTH: 1601
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (40)..(1092)
US-09-629-469A-12810

Query Match 99.7%; Score 1049.8; DB 28; Length 1601;
Best Local Similarity 99.8%; Pred. No. 1.8e-275;
Matches 1051; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATGGGCGACCCGAAAGCGGAGCGGCGGCTGGATCAGGATCAGATCATCTTCA 60

Db 40 ATGGGCGACCCGAAAGCGGAGCGGCTGGATCAGGATCAGATCATCTTCA 99
Qy 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Db 100 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 159
Qy 121 CGGTTTGTTCATCTTTCCAATCCAGTACCCCTGATATTTGGAAAATGTATAACAGGCACAG 180
Db 160 CGGTTTGTTCATCTTTCCAATCCAGTACCCCTGATATTTGGAAAATGTATAACAGGCACAG 219
Qy 181 GCTTCCTCTTGGACAGCAGAAAGGTCGATTTCAAAAGGATCTCCCTCACTGGAACAAG 240
Db 220 GCTTCCTCTTGGACAGCAGAAAGGTCGATTTCAAAAGGATCTCCCTCACTGGAACAAG 279
Qy 241 CTTAAAGCAGATCAGAAAGTACTTCATCTCTCACATCTTAGCCTTTTTCGAGCCAGTGAT 300
Db 280 CTTAAAGCAGATCAGAAAGTACTTCATCTCTCACATCTTAGCCTTTTTCGAGCCAGTGAT 339
Qy 301 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360
Db 340 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 399
Qy 361 CGCTGTTTCTATGCGCTTCAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
Db 400 CGCTGTTTCTATGCGCTTCAAATTTCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 459
Qy 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATGCAA 480
Db 460 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGGAATTTTATTAATGCAATGCAA 519
Qy 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCCCTTGCATGATGATAGAGATAGAAA 540
Db 520 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCCCTTGCATGATGATAGAGATAGAAA 579
Qy 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTAGAGAGGAGTTTCTTCTCAGGA 600
Db 580 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTAGAGAGGAGTTTCTTCTCAGGA 639
Qy 601 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGCTTATGCCAGAGACTCACATTTTTC 660
Db 640 TCTTTTGTGCTATATTTCTGGCTTAAAGAGAGAGGCTTATGCCAGAGACTCACATTTTTC 699
Qy 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
Db 700 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 759
Qy 721 TACTTAGTAAATAAGCCTTCAGAAAGAGGCTCAGGAGAGATCATTTGTTGATGCTGTCAA 780
Db 760 TACTTAGTAAATAAGCCTTCAGAAAGAGGCTCAGGAGAGATCATTTGTTGATGCTGTCAA 819
Qy 781 ATTGAGCAGGAGTTTTTAAACAGAGCCTTGCAGTTGGCCTCATTTGGAATGAATTCATT 840
Db 820 ATTGAGCAGGAGTTTTTAAACAGAGCCTTGCAGTTGGCCTCATTTGGAATGAATTCATT 879
Qy 841 TTGATGAAACAGTACATTTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 900
Db 880 TTGATGAAACAGTACATTTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 939
Qy 901 AAGGTTTTTTCAGCGAGAAAATCCTTTTGTATTTATGAAAAACATTTCTTTAGAAAGGAAAA 960
Db 940 AAGGTTTTTTCAGCGAGAAAATCCTTTTGTATTTATGAAAAACATTTCTTTAGAAAGGAAAA 999
Qy 961 ACAAAATTTCTTTGAGAAAACGAGTTTCAGAGTATCAGCGTTTTCAGAGTTATGGCAGAAACC 1020
Db 1000 ACAAAATTTCTTTGAGAAAACGAGTTTCAGAGTATCAGCGTTTTCAGAGTTATGGCAGAAACC 1059
Qy 1021 ACAGATAACGCTTCTACCTTGGATGCAGATTTT 1053
Db 1060 ACAGATAACGCTTCTACCTTGGATGCAGATTTT 1092

Job time : 3218.5 secs

.....

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 30, 2005, 07:04:13 ; Search time 2982 Seconds
(without alignments)

4805.127 Million cell updates/sec

Title: US-10-698-228-1

Perfect score: 1821

Sequence: 1 MGDPRPBAAGLDQDERSSS.....QRFVMAETTDNVFTLDADF 351

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 45554873 seqs, 20411521753 residues

Total number of hits satisfying chosen parameters: 91109746

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV=xlh

-Q=/cgn2_1/USPTO spoel/US10698228/runat_26102005_100715_5163/app_query.fasta_1.519
-DB=Pending Patents NA Main -QFMT=fastap -SUFFIX=p2n.rnnp -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40 cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptco -NORW=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=200000000 -USR=US10698228 @cgn_1_1_3250 @runat_26102005_100715_5163
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOC=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=120 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Pending Patents_NA_Main:*

- 1: /cgn2_6/ptodata/1/pna/PCTUS1 COMB.seq:*
- 2: /cgn2_6/ptodata/1/pna/PCTUS2 COMB.seq:*
- 3: /cgn2_6/ptodata/1/pna/PCTUS COMB.seq:*
- 4: /cgn2_6/ptodata/1/pna/US06 COMB.seq:*
- 5: /cgn2_6/ptodata/1/pna/US07 COMB.seq:*
- 6: /cgn2_6/ptodata/1/pna/US080 COMB.seq:*
- 7: /cgn2_6/ptodata/1/pna/US081 COMB.seq:*
- 8: /cgn2_6/ptodata/1/pna/US082 COMB.seq:*
- 9: /cgn2_6/ptodata/1/pna/US083 COMB.seq:*
- 10: /cgn2_6/ptodata/1/pna/US084 COMB.seq:*
- 11: /cgn2_6/ptodata/1/pna/US085 COMB.seq:*
- 12: /cgn2_6/ptodata/1/pna/US086 COMB.seq:*
- 13: /cgn2_6/ptodata/1/pna/US087 COMB.seq:*
- 14: /cgn2_6/ptodata/1/pna/US088 COMB.seq:*
- 15: /cgn2_6/ptodata/1/pna/US089 COMB.seq:*
- 16: /cgn2_6/ptodata/1/pna/US090 COMB.seq:*
- 17: /cgn2_6/ptodata/1/pna/US091 COMB.seq:*
- 18: /cgn2_6/ptodata/1/pna/US092A COMB.seq:*
- 19: /cgn2_6/ptodata/1/pna/US092B COMB.seq:*
- 20: /cgn2_6/ptodata/1/pna/US093A COMB.seq:*
- 21: /cgn2_6/ptodata/1/pna/US093B COMB.seq:*
- 22: /cgn2_6/ptodata/1/pna/US094 COMB.seq:*
- 23: /cgn2_6/ptodata/1/pna/US095A COMB.seq:*
- 24: /cgn2_6/ptodata/1/pna/US095B COMB.seq:*
- 25: /cgn2_6/ptodata/1/pna/US095C COMB.seq:*
- 26: /cgn2_6/ptodata/1/pna/US095D COMB.seq:*
- 27: /cgn2_6/ptodata/1/pna/US096A COMB.seq:*
- 28: /cgn2_6/ptodata/1/pna/US096B COMB.seq:*

- 29: /cgn2_6/ptodata/1/pna/US096C COMB.seq:*
- 30: /cgn2_6/ptodata/1/pna/US096D COMB.seq:*
- 31: /cgn2_6/ptodata/1/pna/US096E COMB.seq:*
- 32: /cgn2_6/ptodata/1/pna/US097A COMB.seq:*
- 33: /cgn2_6/ptodata/1/pna/US097B COMB.seq:*
- 34: /cgn2_6/ptodata/1/pna/US097C COMB.seq:*
- 35: /cgn2_6/ptodata/1/pna/US098A COMB.seq:*
- 36: /cgn2_6/ptodata/1/pna/US098B COMB.seq:*
- 37: /cgn2_6/ptodata/1/pna/US098C COMB.seq:*
- 38: /cgn2_6/ptodata/1/pna/US098D COMB.seq:*
- 39: /cgn2_6/ptodata/1/pna/US099A COMB.seq:*
- 40: /cgn2_6/ptodata/1/pna/US099B COMB.seq:*
- 41: /cgn2_6/ptodata/1/pna/US099C COMB.seq:*
- 42: /cgn2_6/ptodata/1/pna/US099D COMB.seq:*
- 43: /cgn2_6/ptodata/1/pna/US099E COMB.seq:*
- 44: /cgn2_6/ptodata/1/pna/US099F COMB.seq:*
- 45: /cgn2_6/ptodata/1/pna/US099G COMB.seq:*
- 46: /cgn2_6/ptodata/1/pna/US100A COMB.seq:*
- 47: /cgn2_6/ptodata/1/pna/US100B COMB.seq:*
- 48: /cgn2_6/ptodata/1/pna/US101A COMB.seq:*
- 49: /cgn2_6/ptodata/1/pna/US101B COMB.seq:*
- 50: /cgn2_6/ptodata/1/pna/US102A COMB.seq:*
- 51: /cgn2_6/ptodata/1/pna/US102B COMB.seq:*
- 52: /cgn2_6/ptodata/1/pna/US103A COMB.seq:*
- 53: /cgn2_6/ptodata/1/pna/US103B COMB.seq:*
- 54: /cgn2_6/ptodata/1/pna/US104A COMB.seq:*
- 55: /cgn2_6/ptodata/1/pna/US104B COMB.seq:*
- 56: /cgn2_6/ptodata/1/pna/US105A COMB.seq:*
- 57: /cgn2_6/ptodata/1/pna/US105B COMB.seq:*
- 58: /cgn2_6/ptodata/1/pna/US106A COMB.seq:*
- 59: /cgn2_6/ptodata/1/pna/US107A COMB.seq:*
- 60: /cgn2_6/ptodata/1/pna/US107B COMB.seq:*
- 61: /cgn2_6/ptodata/1/pna/US107C COMB.seq:*
- 62: /cgn2_6/ptodata/1/pna/US107D COMB.seq:*
- 63: /cgn2_6/ptodata/1/pna/US108A COMB.seq:*
- 64: /cgn2_6/ptodata/1/pna/US108B COMB.seq:*
- 65: /cgn2_6/ptodata/1/pna/US109A COMB.seq:*
- 66: /cgn2_6/ptodata/1/pna/US109B COMB.seq:*
- 67: /cgn2_6/ptodata/1/pna/US109C COMB.seq:*
- 68: /cgn2_6/ptodata/1/pna/US110 COMB.seq:*
- 69: /cgn2_6/ptodata/1/pna/US6000 COMB.seq:*
- 70: /cgn2_6/ptodata/1/pna/US6001 COMB.seq:*
- 71: /cgn2_6/ptodata/1/pna/US6002 COMB.seq:*
- 72: /cgn2_6/ptodata/1/pna/US6003 COMB.seq:*
- 73: /cgn2_6/ptodata/1/pna/US6004 COMB.seq:*
- 74: /cgn2_6/ptodata/1/pna/US6005 COMB.seq:*
- 75: /cgn2_6/ptodata/1/pna/US6006 COMB.seq:*
- 76: /cgn2_6/ptodata/1/pna/US6007 COMB.seq:*
- 77: /cgn2_6/ptodata/1/pna/US6008 COMB.seq:*
- 78: /cgn2_6/ptodata/1/pna/US6009 COMB.seq:*
- 79: /cgn2_6/ptodata/1/pna/US6010 COMB.seq:*
- 80: /cgn2_6/ptodata/1/pna/US6011 COMB.seq:*
- 81: /cgn2_6/ptodata/1/pna/US6012 COMB.seq:*
- 82: /cgn2_6/ptodata/1/pna/US6013 COMB.seq:*
- 83: /cgn2_6/ptodata/1/pna/US6014 COMB.seq:*
- 84: /cgn2_6/ptodata/1/pna/US6015 COMB.seq:*
- 85: /cgn2_6/ptodata/1/pna/US6016 COMB.seq:*
- 86: /cgn2_6/ptodata/1/pna/US6017 COMB.seq:*
- 87: /cgn2_6/ptodata/1/pna/US6018 COMB.seq:*
- 88: /cgn2_6/ptodata/1/pna/US6019 COMB.seq:*
- 89: /cgn2_6/ptodata/1/pna/US6020 COMB.seq:*
- 90: /cgn2_6/ptodata/1/pna/US6021 COMB.seq:*
- 91: /cgn2_6/ptodata/1/pna/US6022 COMB.seq:*
- 92: /cgn2_6/ptodata/1/pna/US6023A COMB.seq:*
- 93: /cgn2_6/ptodata/1/pna/US6023B COMB.seq:*
- 94: /cgn2_6/ptodata/1/pna/US6024 COMB.seq:*
- 95: /cgn2_6/ptodata/1/pna/US6025 COMB.seq:*
- 96: /cgn2_6/ptodata/1/pna/US6026 COMB.seq:*
- 97: /cgn2_6/ptodata/1/pna/US6027 COMB.seq:*
- 98: /cgn2_6/ptodata/1/pna/US6028 COMB.seq:*
- 99: /cgn2_6/ptodata/1/pna/US6029 COMB.seq:*
- 100: /cgn2_6/ptodata/1/pna/US6030 COMB.seq:*
- 101: /cgn2_6/ptodata/1/pna/US6031 COMB.seq:*

Result No.	Query Match	Score	Length	DB ID	Description
102:	/cgn2_6/ptodata/1/pna/US6032	COMB.seq.*			
103:	/cgn2_6/ptodata/1/pna/US6033	COMB.seq.*			
104:	/cgn2_6/ptodata/1/pna/US6034	COMB.seq.*			
105:	/cgn2_6/ptodata/1/pna/US6035	COMB.seq.*			
106:	/cgn2_6/ptodata/1/pna/US6036	COMB.seq.*			
107:	/cgn2_6/ptodata/1/pna/US6037	COMB.seq.*			
108:	/cgn2_6/ptodata/1/pna/US6038	COMB.seq.*			
109:	/cgn2_6/ptodata/1/pna/US6039	COMB.seq.*			
110:	/cgn2_6/ptodata/1/pna/US6040	COMB.seq.*			
111:	/cgn2_6/ptodata/1/pna/US6041	COMB.seq.*			
112:	/cgn2_6/ptodata/1/pna/US6042	COMB.seq.*			
113:	/cgn2_6/ptodata/1/pna/US6043	COMB.seq.*			
114:	/cgn2_6/ptodata/1/pna/US6044	COMB.seq.*			
115:	/cgn2_6/ptodata/1/pna/US6045	COMB.seq.*			
116:	/cgn2_6/ptodata/1/pna/US6046	COMB.seq.*			
117:	/cgn2_6/ptodata/1/pna/US6047	COMB.seq.*			
118:	/cgn2_6/ptodata/1/pna/US6048	COMB.seq.*			
119:	/cgn2_6/ptodata/1/pna/US6049	COMB.seq.*			
120:	/cgn2_6/ptodata/1/pna/US6050	COMB.seq.*			
121:	/cgn2_6/ptodata/1/pna/US6051	COMB.seq.*			
122:	/cgn2_6/ptodata/1/pna/US6052	COMB.seq.*			
123:	/cgn2_6/ptodata/1/pna/US6053	COMB.seq.*			
124:	/cgn2_6/ptodata/1/pna/US6054	COMB.seq.*			
125:	/cgn2_6/ptodata/1/pna/US6055	COMB.seq.*			
126:	/cgn2_6/ptodata/1/pna/US6056	COMB.seq.*			
127:	/cgn2_6/ptodata/1/pna/US6057	COMB.seq.*			
128:	/cgn2_6/ptodata/1/pna/US6058	COMB.seq.*			
129:	/cgn2_6/ptodata/1/pna/US6059	COMB.seq.*			
1	1821	100.0	1053	58	US-10-698-228-2
2	1821	100.0	1053	58	US-10-698-228-12
3	1821	100.0	1081	58	US-10-698-228-4
4	1821	100.0	2596	34	US-09-787-491B-28
5	1821	100.0	2596	81	US-60-128-660-20
6	1821	100.0	3397	49	US-10-170-235-20856
7	1821	100.0	4955	1	PCT-US02-18947-71
8	1821	100.0	4955	2	PCT-US04-01100-71
9	1821	100.0	4955	49	US-10-172-118-71
10	1821	100.0	4955	52	US-10-342-887-71
11	1821	100.0	4955	58	US-10-698-228-3
12	1817	99.8	1601	28	US-09-629-469A-12810
13	1817	99.8	1601	65	US-10-917-503-12810
14	1815	99.7	1849	67	US-10-990-328-49
15	1815	99.7	1849	120	US-60-505-218-20
16	1815	99.7	3397	114	US-60-449-629-12
17	1808	99.3	2597	97	US-60-278-258-266
18	1720.5	94.5	5609	26	US-09-584-852-6398
19	1704	93.6	1140	48	US-10-144-771-870
20	1704	93.6	1140	106	US-60-360-207-870
21	1460.5	80.2	1649	2	PCT-US04-09289-583
22	1460.5	80.2	1989	1	PCT-US00-05882-505
23	1460.5	80.2	1989	40	US-09-925-301-505
24	1460.5	80.2	2215	2	PCT-US03-17409-38
25	1460.5	80.2	2216	47	US-10-084-817-342
26	1460.5	80.2	2216	86	US-60-172-373-422
27	1460.5	80.2	2479	66	US-10-940-774-2025
28	1460.5	80.2	2482	1	PCT-US01-04926A-514
29	1460.5	80.2	2482	23	US-09-515-126-13666
30	1460.5	80.2	2482	26	US-09-577-409-13666
31	1460.5	80.2	2482	50	US-10-220-335-514
32	1460.5	80.2	2482	51	US-10-290-752-514
33	1460.5	80.2	2500	1	PCT-US02-10824-4
34	1460.5	80.2	2500	29	US-09-658-659-9
35	1460.5	80.2	2500	29	US-09-658-659B-9

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of

QY 121 ArgCysPheThrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 DB 361 CGCTGTTTCTATGGCTTTCAAAATTCATCATGAGAAATGTTCACTCAGAGATGTACAGTTG 420
 QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 DB 421 CTGATAGACACTTACATCAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 480
 QY 161 ThrMetProTyrValLysLysValAspTyrAlaLeuArgTyrIleAlaAspArgLys 180
 DB 481 ACCATGCCCTATGTTAAGAAAAAGAGATGGGCTTGGATGGATAGAGATAGAAAA 540
 QY 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
 DB 541 TCTACTTTTGGGGAAGAGTGGTGGCTTGGCTGTAGAGAGATTTCTTCTCAGGA 600
 QY 201 SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 DB 601 TCTTTTGTCTGTATATCTGGCTTAAAGAGAGAGGTCCTTATGCGAGACTCACTTTTTC 660
 QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 DB 661 AATGAACCTATCAGCAGATGAGGACTTCACTGTGACTTTTGTGCTGTGATGTTCCAA 720
 QY 241 TyrLeuValAsnLysProSerGluArgValArgGluIleIleValAspAlaValLys 260
 DB 721 TACTTAGTAAATAGGCTTCAGAGAAAGGTCAGGAGATCATTGTTGATGCTGTCAA 780
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 DB 781 ATTGACGAGGAGTTTTAAACAGAGCCTTGGCCATGTTGGCTCATTTGGAATGCAAT 840
 QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 DB 841 TTGATGAAACAGTACATTCAGTTGTAGCTGACAGATTACTTGTGGAACCTTGGATCTCA 900
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 DB 901 AAGGTTTTTTCAGCAGAAAAATCCTTTGATTTATGGAAGAACATTTCTTTTAGAAGANAA 960
 QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 DB 961 ACAATTTCTTTGAGAAAGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC 1020
 QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 DB 1021 ACAGATAACGTTCTCACCTTGGATGACAGATTTT 1053

RESULT 2

US-10-698-228-12
 ; Sequence 12, Application US/10698228
 ; GENERAL INFORMATION:
 ; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
 ; TITLE OF INVENTION: New Protein and its DNA
 ; FILE REFERENCE: 2619WOOP
 ; CURRENT APPLICATION NUMBER: US/10/698,228
 ; CURRENT FILING DATE: 2003-10-30
 ; PRIOR APPLICATION NUMBER: US/10/019,733
 ; PRIOR FILING DATE: 2001-12-28
 ; PRIOR APPLICATION NUMBER: JP 11-181131
 ; PRIOR FILING DATE: 1999-06-28
 ; PRIOR APPLICATION NUMBER: JP 11-192391
 ; PRIOR FILING DATE: 1999-07-06
 ; PRIOR APPLICATION NUMBER: JP 2000-017770
 ; PRIOR FILING DATE: 2000-01-21
 ; NUMBER OF SEQ ID NOS: 14
 ; SEQ ID NO 12
 ; LENGTH: 1053
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION:

US-10-698-228-12

Alignment Scores:
 Pred. No.: 3,78e-194 Length: 1053
 Score: 1821.00 Matches: 351
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 58 Gaps: 0

US-10-698-228-1 (1-351) x US-10-698-228-12 (1-1053)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20
 DB 1 ATGGCGCACC CGGAAAGCGCGGAGCGCGGCTGGATCAGATCAGATGAGAGATCATCTTCA 60
 QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
 DB 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
 QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrPheLysMetTyrLysGlnAlaGln 60
 DB 121 CGGTTTGTCTATCTTCCAATCCAGTACCTCATATTTGGAAAAATGTATAAACAGGCACAG 180
 QY 61 AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
 DB 181 GCTTCCTTCTGGACAGCAGAGAGGTTGACTTTAAGAGGATCTCCCTCACTGGAAACAG 240
 QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAspSerAsp 100
 DB 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCACAATCTTAGCCTTTTTCAGGCAGTGTAT 300
 QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 DB 301 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGAGTGCAGGTTCCAGAGGCT 360
 QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 DB 361 CGCTGTTTCTATGGCTTTCAAAATTCATCATGAGAAATGTTCACTCAGAGATGTACAGTTG 420
 QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 DB 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATGAA 480
 QY 161 ThrMetProTyrValLysLysLysAlaAspTyrAlaLeuArgTyrIleAlaAspArgLys 180
 DB 481 ACCATGCCCTATGTTAAGAAAAAGAGATGGGCTTGGATGGATAGAGATAGAAAA 540
 QY 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
 DB 541 TCTACTTTTGGGGAAGAGTGGTGGCTTGTGCTGTAGAAAGAGATTTTCTTCTCAGGA 600
 QY 201 SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 DB 601 TCTTTTGTCTGTATATCTGGCTTAAAGAGAGAGGCTTATGCCAGGACTCACTTTTTC 660
 QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 DB 661 AATGAACCTATCAGCAGATGAGGACTTCACTGTGACTTTTGTGCTGTGATGTTCCAA 720
 QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
 DB 721 TACTTAGTAAATAGGCTTCAGAGAAAGGTCAGGAGATCATTGTTGATGCTGTCAA 780
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 DB 781 ATTGACGAGGAGTTTTAAACAGAGCCTTCCAGTGTGCTCATTTGGAATGCAAT 840
 QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 DB 841 TTGATGAAACAGTACATTCAGTTGTAGCTGACAGATTACTTGTGGAACCTTGGATCTCA 900
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320

Db 901 AAGGTTTTTCAGCAGAAATCCTTTGATTTATGGAACAATTTCTTTAGAGGAAAA 960
Qy 321 ThrAsnPhelGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 961 ACAAAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTCAGCTTATGCAGAAACC 1020
Qy 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1021 ACAGATAACGTCCTTACCTTGGATGCAGATTTT 1053

RESULT 3

US-10-698-228-4
; Sequence 4, Application US/10698228
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA
; TITLE OF INVENTION: New Protein and its DNA
; FILE REFERENCE: 2619W00P
; CURRENT APPLICATION NUMBER: US/10/698,228
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 4
; LENGTH: 1081
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-4

Alignment Scores:
Pred. No.: 3,92e-194 Length: 1081
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 58 Gaps: 0

US-10-698-228-1 (1-351) x US-10-698-228-4 (1-1081)

Qy 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
Db 20 ATGGGCGACCCGGAAAGCGCGNAGCGCGGCTGGATCAGATCAGATCATCTTCA 79
Qy 21 AspThrAsnGluSerGluLeuLysSerAsnGluGluProLeuLeuArgLysSerArg 40
Db 80 GACACCAACGAAAGTCAATGAAATGAAAGAGCCCTCTTAAGAAAGAGTTCTCGC 139
Qy 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
Db 140 CGGTTTGTCTCTTCCATTCAGTACCTCCCTGATATTTGGAAATGATATAACAGGCACAG 199
Qy 61 AlaSerPheTrpThrAlaGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
Db 200 GCTTCCTCTTGGACACAGAGAGGTGATCTTCAAGATCTCCCTACCTGGAACAAG 259
Qy 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAlaSerAsp 100
Db 260 CTTAAGCAGATGAGAGTACTTCACTCTCACAATCTTAGCCCTTTTTCAGCCAGTGAT 319
Qy 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnProGluAla 120
Db 320 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 379
Qy 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140

Db 380 CGCTGTTTCTATGGCTTTCAAATTTCTCATCGAGAAATGTTCACTCAGAGATGTTACAGTTTG 439
Qy 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 440 CTGATAGACACTTACATCAGATCCCAAGAAAGGGAATTTTATTATTAATGCAATTGAA 499
Qy 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Db 500 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGCCTTCGATGATAGCAGATAGAAA 559
Qy 181 SerThrPheGlyGluArgValAlaPheAlaValGluGluGlyValPhePheSerGly 200
Db 560 TCTACTTTTGGGAAAAAGAGTGGTGGCTTTGCTGCTAGAGAGGAGTTTCTTCTCAGA 619
Qy 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 620 TCTTTTGGCTGCTATATTTCTGGCTAAAGAAAGAGAGTCTTATGCGCAGGAGTCTTCTTCC 679
Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 680 AATGAATCTCATCAGCAGATGAAGACTTCACCTGCTGCTTGGCTTGCCTGATGTTCCAA 739
Qy 241 TyrLeuValAsnLysProSerGluGluArgValArgGluLeuIleValAspAlaValLys 260
Db 740 TACTTAGTAAATTAAGCCTTCAGAAAGAAAGGTCAGGAGATCATTTGTTGATGCTGCAAA 799
Qy 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 800 ATTGAGCAGGAGTTTAAACAGAGCCTTCCAGCTTGGCCTCATTTGGAATGAATGCAAT 859
Qy 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 860 TTGATGAACAGTACATTTGATTTGATGCTGACAGATTTACTTGTGGAATCTGGATCTCA 919
Qy 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 920 AAGGTTTTTCAGGACAGAAATCCTTTTGAATTTATGGAACAATTTCTTTAGAGGAAAA 979
Qy 321 ThrAsnPhelGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 980 ACAAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGAGTTATGCGCAGAAACC 1039
Qy 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1040 ACAGATAACGTCCTTACCTTGGATGCAGATTTT 1072

RESULT 4

US-09-787-491B-28
; Sequence 28, Application US/09787491B
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; TANG, Y. Tom;
; APPLICANT: CORLEY, Neil C.; GUEGLER, Karl J.;
; APPLICANT: GORGONE, Gina A.; PATTERSON, Chandra;
; APPLICANT: HILLMAN, Jennifer L.; BAUGHN, Mariah R.;
; APPLICANT: LAL, Preeti; AZIMZAI, Valda;
; APPLICANT: YUE, Henry; YANG, Junming
; TITLE OF INVENTION: RNA-ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0600 USN
; CURRENT APPLICATION NUMBER: US/09/787,491B
; CURRENT FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: PCT/US99/21688
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: US 60/128,660
; PRIOR FILING DATE: 1999-04-08
; PRIOR APPLICATION NUMBER: US 60/069,391
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: US 60/183,025
; PRIOR FILING DATE: 1998-09-22
; PRIOR APPLICATION NUMBER: US 60/155,246
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: US 09/158,720
; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 38


```
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 356 CGAATTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT 415

QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 416 CGCTGTTTCTATGGCTTTCAAAATTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 475

QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 476 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAATTTTATTAATGCAATGAA 535

QY 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Db 536 ACCATGCCCTATGTTAAGAAAAGAGATGGGCCCTTGCTGCTGTAGAAGGAGTTTCTTCTCAGGA 655

QY 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
Db 596 TCTACTTTTGGGGAAGAGTGGTGGCCCTTGCTGCTGTAGAAGGAGTTTCTTCTCAGGA 655

QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 656 TCTTTTGTGCTATATCTGGCTTAAGAAGAGAGGCTTATGCCAGAGACTCATTCTTTC 715

QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 716 AATGAACCTCATCAGCAGAGATGAAGACITTCAGTGTGACTTTGCTTGGCTGTATCCAA 775

QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
Db 776 TACTTAGTAATAAAGCCTTCAGAAAGAGGTGAGGAGATCATTTGATGCTGTCAAA 835

QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 836 ATTGAGCAGAGTTTTTAAACAGAGCCTTGCCAGTTGGCTCATTCATGGAATGAATTCAT 895

QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 896 TTGATGAACACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGAACTTGGATCTCA 955

QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 956 AAGGTTTTTTCAGGCAGAAAATCCTTTGATTTTATGAAAACATTTCTTTAGAAGGAAA 1015

QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 1016 ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGATTATGGCAGAAACC 1075

QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1076 ACAGATAACGTTCTACCTTGGNTGCAGATTTT 1108

RESULT 6
US-10-170-235-20856
; Sequence 20856, Application US/10170235
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN
; FILE REFERENCE: CL001380
; CURRENT APPLICATION NUMBER: US/10/170,235
; CURRENT FILING DATE: 2003-03-17
; NUMBER OF SEQ ID NOS: 42514
; SEQ ID NO 20856
; LENGTH: 3397
; TYPE: DNA
; ORGANISM: HUMAN
US-10-170-235-20856

Alignment Scores:
Pred. No.: 1,87e-193 Length: 3397
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
```

```
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 49 Gaps: 0

US-10-698-228-1 (1-351) x US-10-170-235-20856 (1-3397)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20
Db 501 ATGGCGCAGCCCGGAAAGCGCGGAGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 560

QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
Db 561 GACACCAACGAAAGTGAATTAAGTCAATGAAGAGCCACTCTTAAGAGATCTCCCTCACTGGAACAAG 620

QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
Db 621 CGGTTTGTCTATCTTTCCATCCAGTACCTTGATATTTGGAAAATGTATAAACACGACACAG 680

QY 61 AlaSerPheTrpThrAlaGluGluValAspLysSerLysAspLeuProHisTrpAsnLys 80
Db 681 GCCTTCCTTCGGACAGCAGAAAGAGTGCATTTATCAAGGATCTCCCTCACTGGAACAAG 740

QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAspSerAsp 100
Db 741 CTTAAAGCAGATCAGAGTACTTTCTCTCATCTCTAGCTTTTTCAGGAGGAGTAT 800

QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 801 GGAATTTGTAATGAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 860

QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 861 CGCTGTTTCTATGGCTTCAAAATTCATCAGAAATGTTCACTCAGAGATGTACAGTTTG 920

QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 921 CTGATAGACACTTACATCAGAGATCCCAAGAAAAGGAAATTTTATTTAATGCAATTGAA 980

QY 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Db 981 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGCCCTTGGATGGATGAGATAGAGAAA 1040

QY 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
Db 1041 TCTACTTTTGGGGAAGAGTGGTGGCTTTGCTGCTAGAGAGGATTTCTTCTCAGGA 1100

QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 1101 TCTTTTGTCTATATTTGGCTAAAGAGAGAGTCTTATGCCAGGACTCATTCTTTTTC 1160

QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 1161 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTGCTGATGTTTCAA 1220

QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
Db 1221 TACTTAGTAATAAAGCCTTCAGAAAGAGGTGAGGAGATCATTTGTGATGCTGCTCAA 1280

QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 1281 ATTGAGCAGGAGTTTTTAAACAGAAAGCCTTGCCAGTTGGCCTCATTTGGAATGAATTGCATT 1340

QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 1341 TTGATGAACACAGTACATGAGTTTGTAGCTGCACAGATTACTTTGTGGAACATTTCTTCA 1400

QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 1401 AAGTTTTTTCAGGCAGAAAATCCTTTTGAATTTTATGAAAACATTTCTTTTAGAAGAAA 1460

QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 1461 ACAAATTTCTTTGAGAAACGAGTTTCAGAGTATCAGCGTTTTCAGATTATGGCAGAAACC 1520
```


Db 365 CGGTTTGTTCATCTTCCAGTACCCTGATATTTGGAAATGATATAACAGGCACAG 424
Qy 61 AlasexPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
Db 425 GCTTCCTCTTGGACAGAGAGGTCGATTCATCAAGAGATCTCCCTCACTGGAACAAG 484
Qy 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Db 485 CTTAAAGCAGATGAGAAGTACTTCACTCTCAATCTTACGCTTTTGGACGACGATGAT 544
Qy 101 GlyLysValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 545 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 604
Qy 121 ArgCysPheThrGlyPheGlnIleLeuLysValHisSerGluMetTyrSerLeu 140
Db 605 CGCTGTTTCTATGGCTTTCAAAATTCATCGAAGTGTTCACCTCAGAGATGTACAGTTG 664
Qy 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 665 CTGATAGACACTTACATCAGATCCCAAGAAAGGAAATTTTATTTATGCAATGAA 724
Qy 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Db 725 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGCCCTTGGATGATAGCAGATAGAAA 784
Qy 181 SerThrPheGlyGluArgValAlaPheAlaValGluGlyValPhePheSerGly 200
Db 785 TCTACTTTTGGGAAAGAGTGGGCTTTGCTGTGTAAGAGAGTTCCTTCTCAGGA 844
Qy 201 SerPheAlaAlaIlePheThrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 845 TCTTTTGTCTATATCTGGCTTAAGAGAGAGGTCTTATGCGAGGACTCCTTTTTC 904
Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 905 AATGAACTCATCAGCAGATGAAGACTTCCTGTCGCTTGTGCTGCTGATGTTCCAA 964
Qy 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
Db 965 TACTTAGTAATAAGGCTTCAGAAAGAGGCTCAGGAGATCATTTGTGATGCTGTCAA 1024
Qy 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 1025 ATTGACGAGAGTATTTAAACAGAGGCTTGCAGTTGGCCCTCATTTGGAATGATGAT 1084
Qy 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluGlyPheSer 300
Db 1085 TTGATGAAACAGTACATTTGATTTGTAGCTGACAGATTACTTTGGAACTTTGGATTTCA 1144
Qy 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 1145 AAGGTTTTTTCAGGACAGAAAATCCTTTTGAATTTTATGGAACAATTTCTTTAGAGGAAA 1204
Qy 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 1205 ACAAAATTTCTTGAGAAACAGATTCAGATATCAGCGTTTTCAGTTTATGGCAGAAACC 1264
Qy 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1265 ACAGATAACGCTCTTCACTTTGGATGCGATTTT 1297

RESULT 9

US-10-172-118-71

; Sequence 71, Application US/10172118

; GENERAL INFORMATION:

; APPLICANT: Dai, Hongyue

; APPLICANT: He, Yudong

; APPLICANT: Linsley, Peter

; APPLICANT: Mao, Mao

; APPLICANT: Roberts, Chris

; APPLICANT: Van 't Veer, Laura

; APPLICANT: Van de Vijver, Marc

; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 71
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AB036063
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-71

Alignment Scores:

Pred. No.: 3,13e-193 Length: 4955
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 49 Gaps: 0

US-10-698-228-1 (1-351) x US-10-172-118-71 (1-4955)

Qy 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20
Db 245 ATGGCGCACCGAAGCGCGGAGCGCGGCTGGATCAGGATGAGATCATCTTCA 304
Qy 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
Db 305 GACACCAACCAAGTGAATTAAGTCAATGAAGAGCCATCTTAAAGAAAGAGTTCGCG 364
Qy 41 ArgPheValIlePhePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
Db 365 CGGTTTGTCACTTCCATCCAGTACCCTGATATTTGGAAATGTATAAACAGGCACAG 424
Qy 61 AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
Db 425 GCTTCTCTTTCGACAGCAGAGAGGTGCTGACTTATCAAGGATCTCCCTCACTGGAACAAG 484
Qy 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Db 485 CTTAAAGCAGATGAGAAGTACTTCACTCTCAATCTTAGCCTTTTTCAGCCAGTGTAT 544
Qy 101 GlyLysValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 545 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 604
Qy 121 ArgCysPheThrGlyPheGlnIleLeuLysValHisSerGluMetTyrSerLeu 140
Db 605 CGCTGTTTCTATGGCTTTCAAAATTCATCGAAGTGTTCACCTCAGAGATGTACAGTTG 664
Qy 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 665 CTGATAGACACTTACATCAGATCCCAAGAAAGGAAATTTTATTTATGCAATGAA 724
Qy 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Db 725 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGCCCTTGGATGATAGCAGATAGAAA 784
Qy 181 SerThrPheGlyGluArgValAlaPheAlaValGluGlyValPhePheSerGly 200
Db 785 TCTACTTTTGGGAAAGAGTGGGCTTTGCTGTGTAAGAGAGTTCCTTCTCAGGA 844
Qy 201 SerPheAlaAlaIlePheThrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 845 TCTTTTGTCTATATCTGGCTTAAGAGAGAGGTCTTATGCGAGGACTCCTTTTTC 904
Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240

Db 905 AATGAACCTCATCAGCAGAGATGAAGAGCTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 964
 QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleValAspAlaValLys 260
 Db 965 TACTTAGTAATAAGCCTTCAGNAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 1024
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 Db 1025 ATTGAGCAGGAGTATTAACAGAGAGCTTGCAGGCTTGCCTCATTTGGAATGAATGCAAT 1084
 QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuGlyPheSer 300
 Db 1085 TTGATGAACACGTACATGAGTTGTAGCTGCAGATTACTTTGTGAACCTTGGATTTCTCA 1144
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluLys 320
 Db 1145 AAGTTTTTCAGCGAGAAATCTTTTGATTTTATGGAAGAAACATTTCTTTAGAGGAAAA 1204
 QY 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 Db 1205 ACAAAATTTCTTGAGAAACAGTTCAGAGTATCAGAGTATCAGCGTTTGCAGTTATGGCAGAAACC 1264
 QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 Db 1265 ACAGATAACGCTCTTCACTTGGATGCAGATTTT 1297

RESULT 10

US-10-342-887-71

; Sequence 71, Application US/10342887

; GENERAL INFORMATION:

; APPLICANT: Dai, Hongyue

; APPLICANT: He, Yudong

; APPLICANT: Linsley, Peter S.

; APPLICANT: Mao, Mao

; APPLICANT: Roberts, Christopher J.

; APPLICANT: Van 't Veer, Laura Johanna

; APPLICANT: Van de Vijver, Marc J.

; APPLICANT: Bernards, Rene

; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients

; FILE REFERENCE: 9301-188-999

; CURRENT APPLICATION NUMBER: US/10/342,887

; CURRENT FILING DATE: 2003-01-15

; PRIOR APPLICATION NUMBER: 60/298,918

; PRIOR FILING DATE: 2001-06-18

; PRIOR APPLICATION NUMBER: 60/380,710

; PRIOR FILING DATE: 2002-05-14

; PRIOR APPLICATION NUMBER: 10/172,118

; PRIOR FILING DATE: 2002-06-14

; NUMBER OF SEQ ID NOS: 2699

; SEQ ID NO 71

; LENGTH: 4955

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-342-887-71

Alignment Scores:

Pred. No.:	3,136-193	Length:	4955
Score:	1821.00	Matches:	351
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	52	Gaps:	0

US-10-698-228-1 (1-351) x US-10-342-887-71 (1-4955)

QY 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSer 20
 Db 245 ATGGCGAGCCCGGAAGGCCGCGCGGCTGGATCAGGATGAGAGATCATCTTCA 304
 QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
 Db 305 GACACCAAGAAAGTAATAAGTCAATGAGAGCCACTCTTAAGAAGAGTCTCGC 364

RESULT 11

US-10-698-228-3

; Sequence 3, Application US/10698228

; GENERAL INFORMATION:

; APPLICANT: Takeda Chemical Industries, Ltd.; Yusuke NAKAMURA

; TITLE OF INVENTION: New Protein and its DNA

; FILE REFERENCE: 2619W00P

; CURRENT APPLICATION NUMBER: US/10/698,228

; CURRENT FILING DATE: 2003-10-30

QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
 Db 365 CGGTTTGTTCATCTTCCATCCAGTACCTTGATATTTGGAAATGTATAAACAGGCACAG 424
 QY 61 AlaSerPheTrpThrAlaGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
 Db 425 GCTTCTCTTGGCAGCAGAAAGAGTGCATTTATCAAAGGATCTCCCTCACTGGAAACAG 484
 QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAspAsp 100
 Db 485 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTCAGATCTTAGCCTTTTTCAGGCCAGTGA 544
 QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
 Db 545 GGAATTTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGAGAGTGCAGGTTCCAGAGGCT 604
 QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 Db 605 CGCTGTTTCTATGGCTTCAAAATTTCTCATCAGAAATGTTCACTCAGAGATGTACAGTTTG 664
 QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
 Db 665 CTGATAGACACTTACATCAGATGCCAAGAAAGGGAATTTTATTATTAATGCAATTGAA 724
 QY 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
 Db 725 ACCATGCCCTATGTTAAGAAAAAAGCAGATGGGCCCTTGCATGGATGAGTAGAGATAAAA 784
 QY 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
 Db 785 TCTACTTTTGGGGAAGAGTGGTGGCCTTTGCTGCTGTAGAAGGAGTTTCTTCTCAGGA 844
 QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
 Db 845 TCTTTTGTGCTATATTCTGGCTAAAGAGAGAGGCTCTTATGCCAGGAGCTCACATTTTTC 904
 QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 Db 905 AATGAACCTCATCAGCAGAGATGAAGAGCTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 964
 QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
 Db 965 TACTTAGTAATAAGCCTTCAGNAGAAAGGTCAGGAGATCATTTGTTGATGCTGTCAAA 1024
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 Db 1025 ATTGAGCAGGAGTATTAACAGAGAGCTTGCAGTTCGCTCATTTGGAATGAATGCAAT 1084
 QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuGlyPheSer 300
 Db 1085 TTGATGAACACGTACATGAGTTGTAGCTGCAGAGATTACTTTGTGAACCTTGGATTTCTCA 1144
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluLys 320
 Db 1145 AAGTTTTTCAGCGAGAAATCTTTTGATTTTATGGAAGAAACATTTCTTTAGAGGAAAA 1204
 QY 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 Db 1205 ACAAAATTTCTTGAGAAACAGTTCAGAGTATCAGCGTTTGCAGTTATGGCAGAAACC 1264
 QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 Db 1265 ACAGATAACGCTCTTCACTTGGATGCAGATTTT 1297

```

; PRIOR APPLICATION NUMBER: US/10/019,733
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: JP 11-181131
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: JP 11-192391
; PRIOR FILING DATE: 1999-07-06
; PRIOR APPLICATION NUMBER: JP 2000-017770
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 3
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-10-698-228-3

Alignment Scores:
Pred. No.: 3,13e-193 Length: 4955
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 58 Gaps: 0

US-10-698-228-1 (1-351) x US-10-698-228-3 (1-4955)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
DB 245 ATGGCGAGCCCGAAGCGCGAAGCGCGCGCGCTGGATCAGATCAGATCATCTTCA 304
QY 21 AspThrAsnGluSerGluLeuLeuLysSerAsnGluGluProLeuLeuArgLysSerArg 40
DB 305 GACACCAAGCAAGTGAATAAGTCAATGAAGAGCCACTCCTAAGAAAGAGTTCTCCG 364
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
DB 365 CGTTTGTTCATCTTCCATCCAGTACCTGATATTTGGAAATGTATTAACAGGCACAG 424
QY 61 AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
DB 425 GCTTCCTCTCGACAGCAGAGAGGTGCGACTTATCAAGGATCTCCTCACTGGAAACAAG 484
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
DB 485 CTAAAGCAGATGAGAAGTACTTCTATCTCATCTTACATCTTACCTTTTTCGAGCCAGTGT 544
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 545 GGAATGTAAATGAAAATTTGGTGAGCGCTTGTAGTCAGAGGTGCAGGTTCCAGNGGCT 604
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 605 CGCTGTTCTATAGCTTTCAATTTCTATCGAGATGTTTCACTCAGAGATGTACAGTTTG 664
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheIleAsnAlaIleGlu 160
DB 665 CTGATAGACACTTACATCAGATCCCAAGAAAGGGAAATTTTATTTAATGCAATTTGAA 724
QY 161 ThrMetProTyrValLysLysLysAlaAspTyrAlaLeuArgTyrIleAlaAspArgLys 180
DB 725 ACCATGCCCTATGTTAAGAAAAGCAGATTTGGCCCTTCGATGGATAGATAGATAA 784
QY 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
DB 785 TCTACTTTTGGGAAAGAGTGTGGCTTTGCTGTGTAGAGGAGTTTCTTCTCAGGA 844
QY 201 SerPheAlaAlaIlePheTyrLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 845 TCTTTTGTCTATATTTCTGGCTAAAGAGAGAGAGGTCTTATGCGAGGACTCACTTTTCC 904
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240

```

RESULT 12

```

US-09-629-469A-12810
; Sequence 12810, Application US/09629469A
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: SAITO, KAORU
; APPLICANT: YAMAMOTO, JUNICHI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: WAKAMATSU, AI
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: OTSUKI, TETSUJI
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE
; FILE REFERENCE: 084335/0123
; CURRENT APPLICATION NUMBER: US/09/629,469A
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: JP 1999-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 1999-300253
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/183,322
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 19025
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12810
; LENGTH: 1601
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (40)..(1092)
US-09-629-469A-12810

Alignment Scores:
Pred. No.: 1,89e-193 Length: 1601
Score: 1817.00 Matches: 350

```


Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.72%	Mismatches:	0
Query Match:	99.78%	Indels:	0
DB:	28	Gaps:	0
US-10-698-228-1 (1-351) x US-09-629-469A-12810 (1-1601)			
QY	1	MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer	20
DB	40	ATGGGCGACCCGGGAAGCCGGAGCGCGCGGTGGATCAGATCAGAGATCATCTTCA	99
QY	21	AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg	40
DB	100	GACACCAACGAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAGAGTTCTCGC	159
QY	41	ArgPheValIlePheProIleGlnTyrProAspIleTyrIleTyrIleValGlnAlaGln	60
DB	160	CGGTTTGTCTATCTTTCCATCCAGTACCCCTGATATTGGAAAATGTATAACAGGCACAG	219
QY	61	AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrPheAsnLys	80
DB	220	GGTTCTCTCTGACAGCAGAGAGGTGCTGATATCAAGGATCTCCCTCAGTGAACAAG	279
QY	81	LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp	100
DB	280	CTTAAAGCAGATGAGAGTACTTCTCTCATCTTAGCCTTTTTCGACCCAGTGTAT	339
QY	101	GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla	120
DB	340	CGAATTTGTAATGAATAATTTGGTGAGCGCTTTAGTCAGAGGTGCAGGTTCCAGAGGCT	399
QY	121	ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu	140
DB	400	CGCTGCTTCTAGGCTTTCAAATCTCATCGAGAATGTTCACTCAGAGATGACAGTTTG	459
QY	141	LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu	160
DB	460	CTGATAGACACTTATCATCAGATCCCAAGAAGAGGAATTTTATTTAATGCAATTGAA	519
QY	161	ThrMetProTyrValLysLysLysAlaAspTyrAlaLeuArgTyrIleAlaAspArgLys	180
DB	520	ACCATGCCCTATGTTAAGAAAAGCAGATTGGCGCTTCGGATGGATAGACAGATAAAA	579
QY	181	SerThrPheGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly	200
DB	580	TCTACTTTTGGGGAAGAGTGGTGGCTTGTGCTGTAGAAGAGATTCTTCTCAGGA	639
QY	201	SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer	220
DB	640	TCTTTTGTCTATATTTCTGGCTAAGAAGAGAGGTCTTATCCGAGACTCACTTTTTC	699
QY	221	AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln	240
DB	700	AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA	759
QY	241	TyrLeuValAsnLysProSerGluArgValArgGluIleIleValAspAlaValLys	260
DB	760	TACTTAGTAAATAAGCCCTTCAGAAGAAAGGGTTCAGGAGATCATTTGTTGATGCTGTCAA	819
QY	261	IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle	280
DB	820	ATTGACAGAGAGTTTAAACAGAACCTTTCAGTGGCTTGGCTCATTTGGAGTGAATTCAT	879
QY	281	LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer	300
DB	880	TTGATGAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGATTTCA	939
QY	301	LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys	320
DB	940	AAGGTTTTTCAGGCAGCAAAATCTTTTGAATTTATGGAACAATTTCTTTAGAGGAAAA	999
QY	321	ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr	340
DB			

Alignment Scores:
Pred. No.: 1,89e-193 Length: 1601
Score: 1817.00 Matches: 350
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.72% Mismatches: 0
Query Match: 99.78% Indels: 0
DB: 65 Gaps: 0

US-10-698-228-1 (1-351) x US-10-917-503-12810 (1-1601)

QY	1	MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer	20
DB	40	ATGGGCGACCCGGGAAGCCGGAGCGCGGTGGATCAGATCAGAGATCATCTTCA	99
QY	21	AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg	40
DB	100	GACACCAACGAAGTGAATAAAGTCAAAATGAAGAGCCACTCTCTAAGAAGAGTTCTCGC	159
QY	41	ArgPheValIlePheProIleGlnTyrProAspIleTyrIleTyrIleValGlnAlaGln	60
DB	160	CGGTTTGTCTATCTTTCCATCCAGTACCCCTGATATTGGAAAATGTATAACAGGCACAG	219

```
QY 61 AlaSerPheThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
DB 220 GCTTCTCTCTGACAGCAGAGGGTCGACTATCAAGGATCTCCTCACTGGACAG 279
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAsp 100
DB 280 CTTAAAGCAGATGAGAAGTACTTCACTCTCATCTTACATCTTAGCCTTTTTCAGCAGCAGTAT 339
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 340 GGAATTTGTAATAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTCAGAGTTCAGAGGCT 399
QY 121 ArgCysPhePheGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 400 CGCTGCTCTATGGCTTTCAAATTCATCGAGAAATGTCATCGAGATGTACAGTTTG 459
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 460 CTGATAGACACTTACATCAGAGATCCCAAGMAAGGGAATTTTATTTAATGCAATTTGAA 519
QY 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
DB 520 ACCATGCCCTATGTTAAAGAAAAAGCAGATTGGGCCCTTCGATGGATAGCAGATAGAAA 579
QY 181 SerThrPheGlyGluArgValAlaPheAlaValGluGlyValPhePheSerGly 200
DB 580 TCTACTTTTGGGNAAGAGTGGTGGCTTTCGCTGTAGAGGAGTTCCTTCTCAGGA 639
QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 640 TCTTTTGTCTATATTTCTGGCTTAAAGAGAGAGGCTTATGCGCAGACTCACTTTTTC 699
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
DB 700 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTTGTCTTGCCTGTATGTTCCAA 759
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
DB 760 TACTTAGTAAATAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGTTGATGCTGTCAA 819
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
DB 820 ATTGAGCAGAGTTTTAAACAGAAAGCTTGGCAGTTGGCTCATTTGGAGTGAATTCAT 879
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
DB 880 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTAATTTGTGGAACCTTGGATTTCTCA 939
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
DB 940 AAGGTTTTTCAGCAGAAATCCTTTTGATTTTATGGAATAACATTTCTTTAGAAAGAAA 999
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
DB 1000 ACAATTTCTTGAGAAACAGTTTCAGAGTATCAGCGTTTTCAGCTTATGGCAGAAACC 1059
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB 1060 ACAGATAACGCTCTTCACTTGGATGAGATTTT 1092

RESULT 14
US-10-990-328-49
; Sequence 49, Application US/10990328
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 1849
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-990-328-49

Alignment Scores:
Pred. No.: 3 87e-193 Length: 1849
Score: 1815.00 Matches: 350
Percent Similarity: 99.72% Conservative: 0
Best Local Similarity: 99.72% Mismatches: 1
Query Match: 99.67% Indels: 0
DB: 67 Gaps: 0

US-10-698-228-1 (1-351) x US-10-990-328-49 (1-1849)
QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20
DB 85 ATGGCGCACC CGGAAAGCGCGGCGGCTGGATCAGGATCAGGATCAGATCATCTTCA 144
QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
DB 145 GACACCAACCGAAAGTGAATTAAGTCAAAATCAAGAGCCACTCTCTAAGAAAGAGTCTCGC 204
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
DB 205 CGTTTGTCTATCTTCCATTCAGTACCTGATATTTGGAAAATGTATAACAGGCACAG 264
QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
DB 265 GCTTCTCTTGGACAGCAGAAAGAGTGCAGCTTATCAAGAGATCTCCCTCACTGGAACAAG 324
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp 100
DB 325 CTTAAAGCAGATGAGAAGTACTTCACTCTCACATCTTAGCCTTTTTCAGCAGCAGTAT 384
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 385 GGAATTTGTAATAAGAAATTTGGTGGAGCGCTTTAGTCAGAGAGTGCAGGTTCCAGAGCT 444
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 445 CGCTGTTTCTATGCTTCAAAATTCATCAGAAATGTTCACTCAGAGATGTACAGTTTG 504
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 505 CTGATAGACACTTACATTCAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTTGAA 564
QY 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
DB 565 ACCATGCCCTTATGTTAAGAAAAAGCAGATTGGGCCCTTGGCATGGATAGCAGATAGAAA 624
QY 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
DB 625 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGTAGAGAGGATTTTCTTCTCAGA 684
QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 685 TCTTTTGTCTATATTTCTGGCTAAGAGAGAGGCTCTTATGCCAGGACTCACTTTTTC 744
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
DB 745 AATGAACCTCATCAGCAGATGAAGGACTTCACTGTGACTTTGTCTGCTCATGTGTCAA 804
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
DB 805 TACTTAGTAAATAAGCCTTCAGAAAGAGGTCAGGAGATCATTTGTGATGCTGTCAA 864
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
DB 865 ATTGAGCAGGAGTTTTTAAACAGAAAGCCTTGCAGTGGCCCTCATTTGGAATGAATTCATT 924
```

QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 DB 925 TTGATGAACACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTCTCA 984
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 DB 985 AAGGTTTTTCAGGCAGAAATCCTTTTGATTTTATGGAATAACATTTCTTTAGAAAGGAAA 1044
 QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 DB 1045 ACMAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTCAGATTATGGCAGAAACC 1104
 QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 DB 1105 ACAGATAACGCTCTTACCTTGGATGCAGATTTT 1137

RESULT 15

US-60-505-218-20
 ; Sequence 20, Application US/60505218
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; FILE OF INVENTION: CANCER, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001482
 ; CURRENT APPLICATION NUMBER: US/60/505,218
 ; NUMBER OF SEQ ID NOS: 22507
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 20
 ; LENGTH: 1849
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-60-505-218-20

Alignment Scores:

Pred. No.: 3,87e-193 Length: 1849
 Score: 1815.00 Matches: 350
 Percent Similarity: 99.72% Conservative: 0
 Best Local Similarity: 99.72% Mismatches: 1
 Query Match: 99.67% Indels: 0
 DB: 120 Gaps: 0

US-10-698-228-1 (1-351) x US-60-505-218-20 (1-1849)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
 DB 85 ATGGCGACCCGGAAGGCGGAGCGGCGCTGGATCAGGATGAGAGATCATCTTCA 144
 QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
 DB 145 GACACCAAGAAAGTGAATAAGTCAATGAAGAGCCACTCTCAAGAAAGAGTTCTCGC 204
 QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
 DB 205 CGSTTTTGTCATCTTCCATCCAGTACCTGATATTGGAAATGATATAACAGGCACAG 264
 QY 61 AlaSerPheTyrThrAlaGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
 DB 265 GCTTCCTCTGGACAGCAGAGAGGTGCGACTTATCAAGAGATCTCCCTCACTGGAACAAG 324
 QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaSerAsp 100
 DB 325 CTTAAGCAGATGAGAAGTACTTCTCATCTCTACATCTTAGCCCTTTTTCAGCCAGTGAT 384
 QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValProGluAla 120
 DB 385 GGAATTTGTAATGAAAATTTTGTGGAGCGCTTAGTCAGAGGTGCGGTTCCAGAGGCT 444
 QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
 DB 445 CGCTGTTTCTATGGCTTTTCAATTTCTCATCGAGATGTTTCACTCAGAGATGTACAGTTG 504
 QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160

DB 505 CTGATACACACTTACATCWGAGATCCCAAGAAAAGGGAATTTTATTATTAATGCAATGAA 564
 QY 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTTrpIleAlaAspArgLys 180
 DB 565 ACCATGCCCTTATGTTAAGAAAAGAGAGATTTGGGCTTTGCCGATGGATAGCAGATGAAA 624
 QY 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
 DB 625 TCTACTTTTGGGGAAGAGTGGTGGCTTTTCTGCTGTAGAAAGGAGTTTCTTCTCAGGA 684
 QY 201 SerPheAlaIlePheTyrPheLysArgGlyLeuMetProGlyLeuThrPheSer 220
 DB 685 TCTTTTCTGCTATATCTGCTAAAGAGAGAGGTCTTATGCCAGGACTCACTTTTTC 744
 QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
 DB 745 AATGAACCTCATCAGCAGAGATGAGGACTTTCACCTGTGACTTTGCTTGCTGTGATGTTCCAA 804
 QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAlaAlaValLys 260
 DB 805 TACTTAGTAATAAGCCTTTCAGAAAGAGGTTCAGGAGATCATTTGTTGATGCTGTCAA 864
 QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
 DB 865 ATTGAGCAGGAGTTTAAACAGAGAGCTTGGCAGTTGGCCTCATTTGGAATGAATTGCATT 924
 QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
 DB 925 TTGATGAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGGAACCTTGGATTTCTCA 984
 QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
 DB 985 AAGTTTTTTCAGGCAGAAATCCTTTTGATTTTATGGAATAACATTTCTTTAGAAAGAAA 1044
 QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
 DB 1045 ACMAATTTCTTTGAGAAACAGATTTTCAGAGTATCAGCGTTTTCAGATTATGGCAGAAACC 1104
 QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
 DB 1105 ACAGATAACGCTCTTACCTTGGATGCAGATTTT 1137

Search completed: October 30, 2005, 09:45:43

Job time : 2999 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 30, 2005, 07:12:02 ; Search time 1266 Seconds
(without alignments)
3986.300 Million cell updates/sec

Title: US-10-698-228-1
Perfect score: 1821
Sequence: 1 MGDPRPEAAGLDQDERSSS.....QRFVMAETTDNVFTLDADF 351

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2723106 seqs, 7188968421 residues

Total number of hits satisfying chosen parameters: 55446212

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US10698228/runat_26102005_100716_5193/app_query.fasta_1.519
-DB=Pending_Patents NA New -OPMT=fastap -SUFFIX=p2n.rnpn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR_WAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0
-NCPUL=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCk=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:
2: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:
3: /cgn2_6/ptodata/1/pna/US05_NEW_COMB.seq:
4: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq:
5: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq:
6: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:
7: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:
8: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:
9: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:
10: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:
11: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:
12: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:
13: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:
14: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:
15: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:
16: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:
17: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:
18: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:
19: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:
20: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq:
21: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq:
22: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq:
23: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq:
24: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq:
25: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq:
26: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq:
27: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq:
28: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query	Length	DB ID	Description
1	1821	100.0	1056	1	PCT-US05-09639-16	Sequence 16, Appl
2	1821	100.0	1056	27	US-11-088-686-16	Sequence 16, Appl
3	1821	100.0	4929	1	PCT-US05-22501-1545	Sequence 1545, Ap
4	1821	100.0	4955	1	PCT-US05-09639-44	Sequence 44, Appl
5	1821	100.0	4955	1	PCT-US05-21650-13	Sequence 13, Appl
6	1821	100.0	4955	27	US-11-088-686-44	Sequence 44, Appl
7	1817	99.8	1056	1	PCT-US05-09639-18	Sequence 18, Appl
8	1817	99.8	1056	27	US-11-088-686-18	Sequence 18, Appl
9	1816	99.7	1056	1	PCT-US05-09639-20	Sequence 20, Appl
10	1816	99.7	1056	27	US-11-088-686-20	Sequence 20, Appl
11	1815	99.7	1849	15	US-10-990-328A-49	Sequence 49, Appl
12	1813	99.6	1056	1	PCT-US05-09639-22	Sequence 22, Appl
13	1813	99.6	1056	27	US-11-088-686-22	Sequence 22, Appl
14	1812	99.5	1056	1	PCT-US05-09639-24	Sequence 24, Appl
15	1812	99.5	1056	1	PCT-US05-09639-26	Sequence 26, Appl
16	1812	99.5	1056	27	US-11-088-686-24	Sequence 24, Appl
17	1812	99.5	1056	27	US-11-088-686-26	Sequence 26, Appl
18	1811	99.5	1056	1	PCT-US05-09639-28	Sequence 28, Appl
19	1811	99.5	1056	27	US-11-088-686-28	Sequence 28, Appl
20	1460.5	80.2	1170	1	PCT-US05-09639-2	Sequence 2, Appli
21	1460.5	80.2	1170	27	US-11-088-686-2	Sequence 2, Appli
22	1460.5	80.2	1649	9	US-10-553-098-583	Sequence 583, App
23	1460.5	80.2	1649	13	US-10-955-054A-135	Sequence 135, App
24	1460.5	80.2	2470	1	PCT-US05-13260-66	Sequence 66, Appl
25	1460.5	80.2	2470	1	PCT-US05-20840-66	Sequence 66, Appl
26	1460.5	80.2	2470	1	PCT-US05-20840-66	Sequence 66, Appl
27	1460.5	80.2	2470	1	PCT-US05-22501-6487	Sequence 6487, Ap
28	1460.5	80.2	2470	15	US-10-940-774A-2025	Sequence 2025, Ap
29	1460.5	80.2	2482	11	US-10-302-689A-107255	Sequence 107255,
30	1460.5	80.2	2500	1	PCT-US05-09639-43	Sequence 43, Appl
31	1460.5	80.2	2500	1	PCT-US05-20435-12	Sequence 12, Appl
32	1460.5	80.2	2500	1	PCT-US05-21650-12	Sequence 12, Appl
33	1460.5	80.2	2500	1	PCT-US05-27243-78	Sequence 78, Appl
34	1460.5	80.2	2500	14	US-10-960-414-154	Sequence 154, App
35	1460.5	80.2	2500	15	US-10-940-774A-145	Sequence 145, App
36	1460.5	80.2	2500	24	US-11-000-688A-1211	Sequence 1211, Ap
37	1460.5	80.2	2500	24	US-11-150-888-12	Sequence 12, Appl
38	1457.5	80.0	2909	15	US-11-088-686-43	Sequence 43, Appl
39	1457.5	80.0	3616	15	US-10-990-328A-4851	Sequence 4852, Ap
40	1456.5	80.0	1170	1	PCT-US05-09639-4	Sequence 4, Appli
41	1456.5	80.0	1170	27	US-11-088-686-4	Sequence 4, Appli
42	1455.5	79.9	1170	1	PCT-US05-09639-6	Sequence 6, Appli
43	1455.5	79.9	1170	27	US-11-088-686-6	Sequence 6, Appli
44	1452.5	79.8	1170	1	PCT-US05-09639-8	Sequence 8, Appli
45	1452.5	79.8	1170	27	US-11-088-686-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
PCT-US05-09639-16
; Sequence 16, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; PRIOR FILING DATE: 2005-03-23
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16

```
/ LENGTH: 1056
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)...(1056)
PCT-US05-09639-16

Alignment Scores:
Pred. No.: 3e-182 Length: 1056
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-10-698-228-1 (1-351) x PCT-US05-09639-16 (1-1056)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
DB 1 ATGGCGACCCGGAAGCGCGAAGCGCGCGGCTGGATCAGATCAGATCATCTTCA 60
QY 21 AspThrAsnGluSerGluLeuLysSerAsnGluGluProLeuLeuArgLysSerArg 40
DB 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
DB 121 CGGTTTGTCTATCTTCCATCCAGTACCTGATATTTGGAAATGATTAACAGGCACAG 180
QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
DB 181 GCTTCTCTTGGACAGCAGAGAGGTGCGATATCAAGAGATCTCCTCACTGGAACAG 240
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAlaSerAsp 100
DB 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCTCACTCTAGCCCTTTTTCGAGCCAGTGAT 300
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 301 GGAATTTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 360
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGACAGTTTG 420
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 421 CTGATAGACACTTATCATCAGATCCCAAGAAAGGAATTTTATTTAATGCAATTGAA 480
QY 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
DB 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGCATGGATAGCAGATAGAAAA 540
QY 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGluValPhePheSerGly 200
DB 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAAGAGAGTTTCTTCTCAGGA 600
QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 601 TCTTTTGTCTATATCTTGGCTAAAGAGAGAGGTCTTATGTCAGGACTCACTTTTTC 660
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
DB 661 AATGAACCTCATCAGCAGATGAAGACTTCACTGTGACTTTTCTGCTGCTGATGTTCCAA 720
QY 241 TyrLeuValAsnLysProSerGluArgValArgGluIleIleValAspAlaValLys 260
DB 721 TACTTAGTAAATGAAGCTTTCAGAAAGAGGCTCAGGAGATCAATTTGTATGCTGTCAAA 780
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
DB 781 ATTGACAGGAGTTTAAACAGAGGCTTGGCAGTTGGCCTCATTTGAATGAATTCATTT 840
```

```
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
DB 841 TTGATGAAACAGTACATTGAGTTTGTAGCTGACAGATTACTTGTGGAACTTGGATTCTCA 900
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
DB 901 AAGGTTTTTCAGGCAGAAAAATCCTTTTGAATTTATGGAACAATTTCTTTAGAGGAAAA 960
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
DB 961 ACAAAATTTCTTTGAGAAACGAGTTTTCAGAGTATCAGCGTTTTCAGTATATGCGAGAAACC 1020
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB 1021 ACAGATAACGTCCTTCACCTTGGATGACAGATTTT 1053
```

RESULT 2

```
US-11-088-686-16
; Sequence 16, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1056)
US-11-088-686-16
```

```
Alignment Scores:
Pred. No.: 3e-182 Length: 1056
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 27 Gaps: 0
```

US-10-698-228-1 (1-351) x US-11-088-686-16 (1-1056)

```
QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
DB 1 ATGGCGACCCGGAAGCGCGAAGCGCGGCTGGATCAGATGAGATCATCTTCA 60
QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
DB 61 GACACCAACGAAAGTGAATTAAGTCAAAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
DB 121 CGGTTTGTCTATCTTCCATCCAGTACCTGATATTTGGAAAAATGATAAAGCAGCAG 180
QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
DB 181 GCTTCTCTTGGACAGCAGAGAGGTGCGATTAACAAGGATCTCCCTCACTGGAACAG 240
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPheAlaAlaSerAsp 100
DB 241 CTTAAAGCAGATGAGAAGTACTTCACTCTCTCACTCTTAGCCCTTTTTCGAGCCAGTGAT 300
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 301 GGAATTTGTAATGAATAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGGTTCAGAGGCT 360
```

```
QY 121 ArgCysPheTyrGlyPheGlnIleLeuLeuGluAenValHisSerGluMetTyrSerLeu 140
Db 361 CGCTGTTTCTATGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 480
QY 161 ThrMetProTyrValLysLysLysAlaAspTTPAlaLeuArgTyrIleAlaAspArgLys 180
Db 481 ACCATGCCCTATGTTAAGAAAAGACAGATTGGGCTTTCGTGTAGAGAGATTCTTCTCAGGA 540
QY 181 SerThrPheGlyGluArgValAlaPheAlaValGluGlyValPhePheSerGly 200
Db 541 TCTACTTTTGGGAAGAGTGGTGGCTTTTGTGTGTAGAGAGATTCTTCTCAGGA 600
QY 201 SerPheAlaAlaIlePheThrPheLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 601 TCTTTTGTGCTATATCTGGCTAAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGGCTGTATGTCCAA 720
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
Db 721 TACTTAGTAAATAGCCTTCAGAAAGAGGTTCAGGAGATCATTTGTTGATGCTGTCAA 780
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 781 ATTGAGCAGGAGTGTAAACAGAGCCTTGGCAGTTGGCTCATTTGGAATGAATGTCATT 840
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 841 TTGATGAACACAGTACATTCAGTTTGTAGCTGACAGATTACTTTGTGAACTTGGATTCTCA 900
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 901 AAGGTTTTTCAGGCAGAAATCCCTTTGATTTATGGAACAATTCITTTAGAGAA 960
QY 321 ThrAsnPheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 961 ACAAAATTTCTTTGAAACAGGTTTCAGAGTATCAGCGTTTGTGAGTTATGCGAAGAAC 1020
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1021 ACAGATAACGTCTTCACCTTGGATGACAGATTTT 1053
```

RESULT 3

```
PCT-US05-22501-1545
; Sequence 1545, Application PC/TUS0522501
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals
; TITLE OF INVENTION: Determining Cancer-Linked Gene and Therapeutic Targets Using
; TITLE OF INVENTION: Molecular Cytogenetic Methods
; FILE REFERENCE: 689290-249
; CURRENT APPLICATION NUMBER: PCT/US05/22501
; PRIOR FILING DATE: 2005-07-07
; PRIOR APPLICATION NUMBER: 60/581,699
; PRIOR FILING DATE: 2004-06-23
; NUMBER OF SEQ ID NOS: 7840
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1545
; LENGTH: 4929
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: cDNA sequence
PCT-US05-22501-1545
```

Alignment Scores:

```
Pred. No.: 2.13e-181 Length: 4929
Score: 1821.00 Matches: 351
```

```
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-10-698-228-1 (1-351) x PCT-US05-22501-1545 (1-4929)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
Db 245 ATGGGCGACCCGGAAGCCGGAAGCGCGCGGTGGATCAGATCAGATCAGATCATCTTCA 304
QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
Db 305 GACACCAACCAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 364
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
Db 365 CGGTTTGTCTACTTTCCAATCCAGTACCTGATATTTGGAATAATGTAATAACAGGCACAG 424
QY 61 AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
Db 425 GCTTCCTTCTGGACAGCAGAGAGGTGCATTTCAAGAGGATCTCCCTCAGTGGAAACAG 484
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Db 485 CTTAAGCAGATGAGAAGTACTTTCATCTCTCACATCTTAGCCTTTTTCAGCCAGTGTAT 544
QY 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 545 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGAGGTGCAGGTTCCAGAGGCT 604
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 605 CGCTGTTTCTATGGCTTTCAAAATTCATCAGAAATGTTCTACTCAGAGATGTACAGTTG 664
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 724
QY 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTyrIleAlaAspArgLys 180
Db 725 ACCATGCCCTATGTTAAGAAAAAAGCAGATTTGGGCTTTCGATGATGATAGCAGATAAAA 784
QY 181 SerThrPheGlyGluArgValAlaPheAlaValGluGlyValPhePheSerGly 200
Db 785 TCTACTTTTGGGAAAAGAGTGGTGGCTTTCCTCTGTAGAAGGAGTTTCTTCTCAGGA 844
QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 845 TCTTTTGTGCTATATTTCTGGCTAAAGAGAGAGTCTTATGCGCAGGACTCACTTTTTC 904
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 905 AATGAACCTCATCAGCAGAGATGAAGACTTTCACCTGTGACTTTTGTCTGCTGATGTTCAA 964
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
Db 965 TACTTAGTAAATAAGCCTTCAGAAGAAAGGCTCAGGAGATCATTTGTTGATGCTGTCAA 1024
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 1025 ATTGAGCAGGAGTTTAAACAGAGAGCCTTGCAGTTGGCCTCATTTGGAATGAATTCATT 1084
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 1085 TTGATGAACACAGTACATTTGATTTGTAGCTGACAGATTACTTTGTGGAACCTTGGATTCTCA 1144
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 1145 AAGGTTTTTTCAGGCAGAAAAATCCTTTTGAATTTATGGAACAATTTCTTTTAGAGGAAA 1204
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
```

Db 1205 ACAAATTTCTTTGAGAAACAGCTTTTCAGAGTATCAGCGTTTTTGCAGTTATATGCGAGAAACC 1264
Qy 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1265 ACAGATAAGGCTTTCACCTTGGATGAGATTTT 1297

RESULT 4

PCT-US05-09639-44
; Sequence 44, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US05-09639-44

Alignment Scores:
Pred. No.: 2,14e-181 Length: 4955
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-10-698-228-1 (1-351) x PCT-US05-09639-44 (1-4955)

Qy 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSer 20
Db 245 ATGGCGAGCCCGAAGGCGGAGCGCGGCTGGATCAGATCAGAGATCATCTTCA 304
Qy 21 AspThrAsnGluSerGluLeuIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
Db 305 GACACCAACGAAAGTGAATAAGTCAATGAAAGAGCCACTCTTAAGAAAGAGTTCTCGC 364
Qy 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
Db 365 CGTTTGTTCATCTTCCATCCAGTACCTCTGATATTTGGAAATGTATTAACAGGCACAG 424
Qy 61 AlaSerPheTyrThrAlaGluGluValAspLeuSerLysAspLeuProHisTyrAsnLys 80
Db 425 GCTTCTCTTCGACAGCAGAGAGTTCGACTTATCAAGAGATCTCCCTCACTGGAAACAAG 484
Qy 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Db 485 CTTAAAGCAGATGAGAAGTACTTCATCTCATCTCTAGCCCTTTTTCAGCCAGTAT 544
Qy 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 545 GGAATGTAAATGAAATTTGTGGAGCGCTTTAGTCAGAGGTGAGGTTCCAGAGGCT 604
Qy 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 605 CGCTGTTCTATGGCTTTCAATCTCATCGAGATGTTCACTCAGAGATGTACAGTTTG 664
Qy 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 665 CTGATAGACACTTACATCAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 724
Qy 161 ThrMetProTyrValLysLysAlaAspTyrAlaLeuArgTyrIleAlaAspArgLys 180
Db 725 ACCATGCCCTATGTTAAGAAAAAGCAGATGGGCTTCCGATGGATAGCAGATAGAAA 784
Qy 181 SerThrPheGlyGluArgValAlaPheAlaAlaValGluGlyValPhePheSerGly 200

Db 785 TCTACTTTTGGGAAACAGGTGGCGCTTGTCTGTCTGTAAGAGGAGTTTTCTTCTCAGA 844
Qy 201 SerPheAlaAlaIlePheTyrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 845 TCTTTTGTCTATATTTCTGGCTTAAGAAAGAGAGTCTTATGCGAGACTCACATTTTCC 904
Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 905 AATGAACCTCATCAGCAGAGATGAAGGACTTCTCTGTCTGCTTGTCTGCTGATGTTCAA 964
Qy 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
Db 965 TACTTGTAAATAAGCCTTCAGAAAGAAAGGTCAGGAGATCATTTGTGTGCTGTCAAA 1024
Qy 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 1025 ATTGACGAGGAGTTTTTAACAGAACCTTGCAGTTGGCTTCATTTGGAATGATTTGCA 1084
Qy 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuGlyPheSer 300
Db 1085 TTGATGAACAGTACATTTGATTTGTAGCTGACAGATTACTTTGTGGAACCTTGGATTCTCA 1144
Qy 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGlyLys 320
Db 1145 AAGGTTTTTTCAGGCAGAAAAATCCTTTTGTATTTATGGAACATTTCTTTAAGAGGAAA 1204
Qy 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 1205 ACAATTTCTTTGAGAAACAGTTTTCAGAGTATCAGCGTTTTTGCAGTTATGCGAGAAACC 1264
Qy 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1265 ACAGATAAGCTTCTTCACTTGGATGACAGATTTT 1297

RESULT 5

PCT-US05-21650-13
; Sequence 13, Application PC/TUS0521650
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: MIGFS AS MODIFIERS OF THE IGF PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX05-017
; CURRENT APPLICATION NUMBER: PCT/US05/21650
; CURRENT FILING DATE: 2005-06-27
; PRIOR APPLICATION NUMBER: 60/581,689
; PRIOR FILING DATE: 2004-06-21
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US05-21650-13

Alignment Scores:
Pred. No.: 2,14e-181 Length: 4955
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-10-698-228-1 (1-351) x PCT-US05-21650-13 (1-4955)

Qy 1 MetGlyAspProGluArgProGluAlaGlyLeuAspGlnAspGluArgSerSer 20
Db 245 ATGGCGAGCCCGAAGGCGGAGCGCGGCTGGATCAGATCAGAGATCATCTTCA 304
Qy 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
Db 305 GACACCAACGAAAGTGAATAAGTCAATGAAAGAGCCACTCTTAAGAAAGAGTTCTCGC 364
Qy 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60

```
Db 365 CGTTTGTCTATCTTCCATCCAGTACCTGATATTTGGAAATGTATAAACAGGCACAG 424
Qy 61 AlaSerPheThrAlaGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
Db 425 GCTTCTTCTGACACAGAGAGGTGCTATCAAGGATCTCTCTCCTCAGTGGAAACAG 484
Qy 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAspAsp 100
Db 485 CTTAAAGCAGATGAGAGTACTTCTCATCTCTAGCCCTTTTGGCCAGCTGAT 544
Qy 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 545 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGTTCCAGAGCT 604
Qy 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 605 CGCTGTTCTATGCGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGACAGTTTG 664
Qy 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 724
Qy 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Db 725 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGCGCTTGGGATGCGATAGCAGATAGAAA 784
Qy 181 SerThrPheGluValArgValAlaPheAlaIleValGluGlyValPhePheSerGly 200
Db 785 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTGTAAGAGGATTTTCTTCTCAGGA 844
Qy 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 845 TCTTTTGTCTATATTTCTGGCTAAAGAAAGAGGTCCTTATGCCAGGACTCACTTTTTC 904
Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 905 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA 964
Qy 241 TyrLeuValAsnLysProSerGluValArgValArgGluIleIleValAspAlaValLys 260
Db 965 TACTTAGTAAATTAAGCCCTTCAGAAAGAAAGGTCAGGAGATCATTTGTTGATCTGTCAA 1024
Qy 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 1025 ATTGACGAGGAGTGTATTAACAGAAAGCCTTGCAGTTGGCCTCATTTGGAATGAATGCA 1084
Qy 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 1085 TTGATGAACAGTACATTTGAGTTTGTAGCTGACAGATTAATTTGTTGGAATCTGATTTCA 1144
Qy 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 1145 AAGGTTTTTTCAGGCAGAAAAATCCTTTGATTTATGGAATAACATTTCTTTAGAGGAAA 1204
Qy 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 1205 ACAAAATTTCTTGAGAAACAGATTTTCAGAGTATCAGCGTTTGGCAGTTATGCGAAGAAC 1264
Qy 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1265 ACAGATAACGCTTTCACCTTGGATGACAGATTTT 1297
```

RESULT 6

```
US-11-088-686-44
; Sequence 44, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
```

```
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 4955
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-088-686-44
```

```
Alignment Scores: 2.14e-181 Length: 4955
Score: 1821.00 Matches: 351
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 27 Gaps: 0
```

US-10-698-228-1 (1-351) x US-11-088-686-44 (1-4955)

```
Qy 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
Db 245 ATGGGCGACCCGGAAGGCCGGAAGCGCGCGCTGGATCAGGATGAGAGATCATTTCA 304
Qy 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
Db 305 GACACCAACGAAAGTGAATATAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 364
Qy 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
Db 365 CGGTTTGTCTATCTTCCAATCCAGTACCTTGATATTGGAAAAATGTATAAACACGCGCAG 424
Qy 61 AlaSerPheTrpThrAlaGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
Db 425 GCTTCTCTTGGACAGCAGAGAGGTCGATTAACAAGGATCTCCCTCACTGGAAACAG 484
Qy 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAspSerAsp 100
Db 485 CTTAAAGCAGATGAGAAGTACTTCACTCTCTCACATCTTAGCCTTTTTCAGCCAGCTGAT 544
Qy 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 545 GGAATTTGTAATGAAATTTGGTGGAGCGCTTTAGTCAGAGGTGCGAGTTCCAGAGGCT 604
Qy 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 605 CGCTGTTTCTATGCGCTTTCAAAATTTCAATCAGAAATGTTCACTCAGAGATGTACAGTTTG 664
Qy 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 665 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTTAATGCAATTGAA 724
Qy 161 ThrMetProTyrValLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Db 725 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGCGCTTGGATGCGATAGCAGATAGAAA 784
Qy 181 SerThrPheGluValArgValAlaPheAlaIleValGluGlyValPhePheSerGly 200
Db 785 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGTCTAGAAAGGAGTTTCTTCTCAGGA 844
Qy 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 845 TCTTTTGTCTATATTTCTGGCTAAAGAGAGGTCCTTATGCCAGGACTCACTTTTTC 904
Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 905 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCAA 964
Qy 241 TyrLeuValAsnLysProSerGluValArgValArgGluIleIleValAspAlaValLys 260
Db 965 TACTTAGTAAATTAAGCCCTTCAGAAAGAAAGGTCAGGAGATCATTTGTTGATCTGTCAA 1024
Qy 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 1025 ATTGACGAGGAGTGTATTAACAGAAAGCCTTGCAGTTGGCCTCATTTGGAATGAATGCA 1084
```


281	LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer	300
	TTGATGAAACAGGTACATTTGAGTTTGTAGCTGCACAGATTACTCTGTGGAACTTGATTGTTCTCA	1144
1085		
301	LysValPheGlnAlaGluAsnProPheAspPheMetClnAsnIleSerLeuGluGlyLys	320
1145	AAGGTTTTTCAGGCAGAAAAATCTTTTGTATTATGTGAAACATTTCTTTAGAAAGGAAAA	1204
321	ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr	340
1205	ACAAATTTCTTCAGAAAAACGAGTTTCAGAGTATCAGCGTTTTCAGCTTATGCGAGAAACC	1264
341	ThrAspAsnValPheThrLeuAspAlaAspPhe	351
1265	ACAGATAACGCTTTCACCTTGGATGCAGATTTT	1297

```

RESULT 7
PCT-US05-09639-18
; Sequence 18, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US05-09639-18

```

Alignment Scores:		
Pred. No.:	7.97e-182	Length: 1056
Score:	1817.00	Matches: 350
Percent Similarity:	100.00%	Conservative: 1
Best Local Similarity:	99.72%	Mismatches: 0
Query Match:	99.78%	Indels: 0
DB:	1	Gaps: 0

US-10-698-228-1 (1-351) x PCT-US05-09639-18 (1-1056)

Qy	1	MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer	20
Db	1	ATGGGGACCCGGAAAGCCGGAAAGCCGGCTGGATCAGGATGAGAGATCATCTTCA	60
Qy	21	AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg	40
Db	61	GACCAACCAAGAGTGAAATAAAGTCAATGAGAGCCACCTCTTAAGAAAGAGTTCTCGC	120
Qy	41	ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln	60
Db	121	CGGTTTGTCATCTTCCAAATCCAGTACCCCTGATATTGGAAAAATGATAAACAGGCACAG	180
Qy	61	AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys	80
Db	181	GCATTCTCTTCGGACAGCAGAGAGGTGCACATTATCAAAAGGATCTCCCTCACTGGAAACAAG	240
Qy	81	LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAspAsp	100
Db	241	CTTAAAGCAGATGAGAAGTACTTCATCTCTCACATCTTAGCGCTTTTTCGAGCCAGTGGAT	300
Qy	101	GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla	120
Db	301	GGAATTGTAATGAAATTTGGTCGAGCGCTTTTAGTCAGGAGGTGCAGGTTCCAGAGCT	360
Qy	121	ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu	140
Db	361	CGCTGTATCTTTGGCTTTTCAAATCTCATCAGAAATGTTCACTCAGAGATGTACAGTTTG	420

Qy		141	LeuIleAspThrTyrTrileArgaspProLysLysArgGluPheAsnAlaIleGlu	160
Db		421	CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAAATTTTATTATGTGAATGAA	480
Qy		161	ThrMetProFtyrValLysLysLysAlaAspTpAlaLeuArgTrrpileAlaAspAqLys	180
Db		481	ACCATGCCCTATGTTTAAGAANAANAAGCAGATTGGGCCCTTGCGATGGATAGCAGATAGAAA	540
Qy		181	SerThrPheGlygluArgValValAlaPheAlaAlaValIcIuGlyValPhePheSerGly	200
Db		541	TCTACTTTTTGGGGAAAGAGTGGTGGCCTTGCTGCTGTAGAAAGAGCTTTCTTCTCAGGA	600
Qy		201	SerPheAlaAlaIlePheTrpleuLysLysArgGlyLeuMetProGlyLeuThrPheSer	220
Db		601	TCFTTTTGCCTATATTCCTGGCTTAAAGAGAGAGGCTTTATGCCAGAGCTCACTTTTTC	660
Qy		221	AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln	240
Db		661	AATGAACCTCATCAGCAGAGATGNAGGACCTTCACTGTGACTTTGCTTGCCTGATGTTCCAA	720
Qy		241	TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys	260
Db		721	TACTTTAGTAAATAAGCGCTTCAGAAGAAGAGGTCAGGAGATCATTTGTGATGCTGTCAA	780
Qy		261	IleGluGluGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle	280
Db		781	ATTGACAGAGAGTTTTTAAACAGNAGCGCTTGCCAGTTGGCCCTCATTTGGAATGAATTGCATT	840
Qy		281	LeuMetLysGlnTyrrileGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer	300
Db		841	TTGATGAACAACAGTACATTTCAGTTTTGTAGCTGCAGAGATTACTTGTGGAACTTGAGATTCTCA	900
Qy		301	LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys	320
Db		901	AAGGTITTTTTCAGGCAGAAAAATCCCTTTTGATTTTATGAAAACAATTTCTTTTAGAAGAGAAA	960
Qy		321	ThrAsnPhePheGluLysArgValSerGluTyrrGlnArgPheAlaValMetalAlaGluThr	340
Db		961	ACAAAATTTCTTTGAGAAAACGAGTTTCAGAGTATCAGCGTTTTCAGTTATGGCAGAAACC	1020
Qy		341	ThrAspAsnValPheThrLeuAspAlaAspPhe	351
Db		1021	ACAGATAAGCTCTTTCACCTTGAATGCAGATTTT	1053

RESULT 8

```

RESULTS
US-11-088-686-18
; Sequence 18, Application US/11088686
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004001
; CURRENT APPLICATION NUMBER: US/11/088,686
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-088-686-18

```

Alignment Scores:		
Pred. No.:	7.97e-182	Length: 1056
Score:	1817.00	Matches: 350
Percent Similarity:	100.00%	Conservative: 1
Best Local Similarity:	99.72%	Mismatches: 0
Query Match:	99.78%	Indels: 0
DB:	27	Gaps: 0

US-10-698-228-1 (1-351) x US-11-088-686-18 (1-1056)

```
Qy 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
Db 1 ATGGCGAGCCCGAAGAGCCGAGCGCGCGCTGGATCAGATGAGATCATCTTCA 60

Qy 21 AspThrAsnGluSerGluLeuLysSerAsnGluGluProLeuLeuArgLysSerArg 40
Db 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAGAGTTCTCGC 120

Qy 41 ArgPheValIlePheProIleGlnTyPProAspIleTTrpLysMetTyriysGlnAlaGln 60
Db 121 CGGTTTGTTCATCTTTCCAAATCAGTACCTGTATATTGGAAATGTTATTAACAGGCACAG 180

Qy 61 AlaSerPheTTrpAlaGluGluValAspLeuSerLysAspLeuProHisTTrpAsnLys 80
Db 181 GCTTCTCTCTGGACAGCAGAGAGGTGCATTTATCAAGGATCTCCCTCAGTGGACAG 240

Qy 81 LeuLysAlaAspGluLysTyPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Db 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTAGCCTTTTTCAGCCAGTGTAT 300

Qy 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 301 GGAATTTGTAATGAAATTTTGGTGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360

Qy 121 ArgCysPheTyGlyPheGlnIleLeuLysValHisSerGluMetTyriSerLeu 140
Db 361 CGCTGTTTCTTTGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420

Qy 141 LeuIleAspThrTyIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 421 CTGATAGACACTTACATGAGATCCCAAGAAAGGGAAATTTTATTAATGCAATTGAA 480

Qy 161 ThrMetProTyValLysLysAlaAspTTrpAlaLeuArgTTrpIleAlaAspArgLys 180
Db 481 ACCATGCCCTATGTTTAAGAAAGAGCAGATTGGCGCTTCGATGGATAGCAGATAGAAAA 540

Qy 181 SerThrPheGlyGluArgValAlaPheAlaValAlaValGluGlyValPhePheSerGly 200
Db 541 TCTACTTTTGGGGAAGAGTGGTGCCCTTTCGCTGTAGAGAGGATTTCTTCTCAGGA 600

Qy 201 SerPheAlaAlaIlePheTTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 601 TCTTTTGTCTATATTCTGCTTAAGAGAGAGAGGTCTTATGCCAGGACTCACTTTTCC 660

Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720

Qy 241 TyrIleValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
Db 721 TACTTAGTAAATAAGCCTTCAGAGAAAGGGTCAGGAGATCATGTTGATGCTGTCAAA 780

Qy 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 781 ATTGAGCAGGAGTTTTTHACAGAGCCCTTGCAGTTGGCCTCATTTGGAAATGAATTGCATT 840

Qy 281 LeuMetLysGlnTyIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 841 TTGATGAACACAGTACATTGAGTTGTAGCTGACAGATTACTTGTGGAACCTTGATTTCA 900

Qy 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 901 AAGGTTTTTTCAGGCAGAGAAATCCCTTTGATTTATGGAAGAACATTTCTTTAGAGGAAAA 960

Qy 321 ThrAsnPhePheGluLysArgValSerGluTyGlnArgPheAlaValMetAlaGluThr 340
Db 961 ACAAAATTTCTTTGAGAAACAGAGTTTCAGAGATATCAGCGTTTTCAGAGTTATGGCAGAAC 1020

Qy 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1021 ACAGATAACGCTTTCACCTTCGATGCAGATTTT 1053
```

RESULT 9

```
PCT-US05-09639-20
; Sequence 20, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004W01
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US05-09639-20
```

Alignment Scores:

Pred. No.:	1,02e-181	Length:	1056
Score:	1816.00	Matches:	350
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.72%	Mismatches:	0
Query Match:	99.73%	Indels:	0
DB:	1	Gaps:	0

US-10-698-228-1 (1-351) x PCT-US05-09639-20 (1-1056)

```
Qy 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
Db 1 ATGGCGAGCCCGAAGAGCCGAGCGCGCGCTGGATCAGATGAGATCATCTTCA 60

Qy 21 AspThrAsnGluSerGluLeuLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
Db 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTCTAAGAAGAGTTCTCGC 120

Qy 41 ArgPheValIlePheProIleGlnTyPProAspIleTTrpLysMetTyriysGlnAlaGln 60
Db 121 CGGTTTGTTCATCTTTCCAAATCAGTACCTCTCATCTCTAGCCTTTTTCAGCCAGTGTAT 180

Qy 61 AlaSerPheTTrpAlaGluGluValAspLeuSerLysAspLeuProHisTTrpAsnLys 80
Db 181 GCTTCTCTCTGGACAGCAGAGAGGTGCATTTATCAAGGATCTCCCTCAGTGGACAG 240

Qy 81 LeuLysAlaAspGluLysTyPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Db 241 CTTAAAGCAGATGAGAAGTACTTCTCATCTCTAGCCTTTTTCAGCCAGTGTAT 300

Qy 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 301 GGAATTTGTAATGAAATTTTGGTGAGCGCTTTAGTCAGGAGGTGCAGGTTCCAGAGGCT 360

Qy 121 ArgCysPheTyGlyPheGlnIleLeuLysValHisSerGluMetTyriSerLeu 140
Db 361 CGCTGTTTCTTTGGCTTTCAAAATCTCATCGAGAATGTTCACTCAGAGATGTACAGTTTG 420

Qy 141 LeuIleAspThrTyIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 421 CTGATAGACACTTACATGAGATCCCAAGAAAGGGAAATTTTATTAATGCAATTGAA 480

Qy 161 ThrMetProTyValLysLysAlaAspTTrpAlaLeuArgTTrpIleAlaAspArgLys 180
Db 481 ACCATGCCCTATGTTTAAGAAAGAGCAGATTGGCGCTTCGATGGATAGCAGATAGAAAA 540

Qy 181 SerThrPheGlyGluArgValAlaPheAlaValAlaValGluGlyValPhePheSerGly 200
Db 541 TCTACTTTTGGGGAAGAGTGGTGCCCTTTCGCTGTAGAGAGGATTTCTTCTCAGGA 600

Qy 201 SerPheAlaAlaIlePheTTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 601 TCTTTTGTCTATATTCTGCTTAAGAGAGAGAGGTCTTATGCCAGGACTCACTTTTCC 660
```


US-10-990-328A-49

Alignment Scores:

Pred. No.: 2,648-181 Length: 1849
Score: 1815.00 Matches: 350
Percent Similarity: 99.72% Conservative: 0
Best Local Similarity: 99.72% Mismatches: 1
Query Match: 99.67% Indels: 0
DB: 15 Gaps: 0

US-10-698-228-1 (1-351) x US-10-990-328A-49 (1-1849)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20
DB 85 ATGGCGACCCGGAAGCGCGAGCGCGCGCTGGATCAGATGAGATCATCTTCA 144
QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40
DB 145 GACACCAACGAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 204
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysMetTyrLysGlnAlaGln 60
DB 205 CGTTTGTTCATCTTCCATCCAGTACCCTGATATTGGAAATGTATAAACAGGCACAG 264
QY 61 AlaSerPheThrAlaGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
DB 265 GTTCTCTCTGACAGCAGAGAGGTGCTTATCAAGGATCTCCTCAGTGGAAACAG 324
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
DB 325 CTAAAGCAGATGAGAGTACTTCTCATCTCAGTCTTGTAGCTTTTTCAGCCAGTAT 384
QY 101 GlyIleValAsnGlnAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 385 CGAATTTGTAATGAAATTTTGTGGAGCGCTTATAGTCAGGAGGTGAGGTTCCAGAGGT 444
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 445 CGCTGTTCTATGGCTTTCAAAATCTCATCGAAGATGTTCTCATGAGATGTACAGTTG 504
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 505 CTGATAGACACTTACATCWAGATCCCAAGAAAGGGAATTTTATTAATGCAATGAA 564
QY 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
DB 565 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGGCTTGGCATGGATAGCAGATAGAAAA 624
QY 181 SerThrPheGluArgValAlaAlaPheAlaValGluGlyValPhePheSerGly 200
DB 625 TCTACTTTTGGGGAAGAGTGGTGCCCTTGTCTGTAGAGAGGATTTTCTTCTCAGGA 684
QY 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
DB 685 TCTTTTGTCTATATTCTGGCTAAGAGAGAGGTCTTATGCCAGGACTCATTCTTCC 744
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
DB 745 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTCCAA 804
QY 241 TyrLeuValAsnLysProSerGluArgValArgGluIleIleValAspAlaValLys 260
DB 805 TACTTAGTAAATTAAGCCTTTCAGAAAGAGGGTCAAGGAGATCATTTGTTGATCTGTCAA 864
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
DB 865 ATTGACGAGAGGATTTTAAACAGAACCTTGCAGTTGGCTCATTTGGAATGATTCAT 924
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuValGluLeuGlyPheSer 300
DB 925 TTGATGAAACAGTACATTTGAGTTGTAGCTGACAGATTACTTTGTGGAACCTTGGATTC 984
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320

DB 985 AAGGTTTTTCAGGCAGAAAAATCCTTTTGTGATTTTATGAAAAACATTTCTTTAGAACGAAAA 1044
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
DB 1045 ACAAAATTTCTTTTGAGAAACAGTTTTCAGAGTATCAGCGTTTTCAGAGTTATGCGAGAAACC 1104
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
DB 1105 ACAGATAACGCTCTTCACCTTGGATGCAGATTTT 1137

RESULT 12

PCT-US05-09639-22
; Sequence 22, Application PC/TUS0509639
; GENERAL INFORMATION:
; APPLICANT: Yen, Yun
; TITLE OF INVENTION: DRUG SCREENING
; FILE REFERENCE: 14037-004WO1
; CURRENT APPLICATION NUMBER: PCT/US05/09639
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/556,836
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US05-09639-22

Alignment Scores:
Pred. No.: 2,116-181 Length: 1056
Score: 1813.00 Matches: 349
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.43% Mismatches: 0
Query Match: 99.56% Indels: 0
DB: 1 Gaps: 0

US-10-698-228-1 (1-351) x PCT-US05-09639-22 (1-1056)

QY 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSerSer 20
DB 1 ATGGCGACCCGGAAGCGCGAGCGCGCTGGATCAGATGAGATCATCTTCA 60
QY 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg 40
DB 61 GACACCAACGAAGTGAATAAAGTCAATGAAGAGCCACTCTCTAAGAAAGAGTTCTCGC 120
QY 41 ArgPheValIlePheProIleGlnTyrProAspIleTyrLysGlnAlaGln 60
DB 121 CGTTTGTCTATCTTCCATCCAGTACCCTGATATTTGGAAAAATGTATAAACAGGCACAG 180
QY 61 AlaSerPheThrAlaGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
DB 181 GCTTCTCTTGGACAGCAGAGAGGTGCACTTATCAAGAGATCTCCCTCATCTGGACACAG 240
QY 81 LeuLysAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
DB 241 CTAAAGCAGATGAGAGTACTTCTCATCTCAGATCTTAGCCCTTTTTCAGCCAGTGCAT 300
QY 101 GlyIleValAsnGlnAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
DB 301 GGAATTTGTAATGAAATTTTGGTGGAGCGCTTTTAGTCAGAGAGGTGCGAGTTCCAGAGCT 360
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
DB 361 CGCTGTTTCTTGGCTTTCAAAATCTCATCGAAGATGTTCTCATGAGATGTTTAGTTTG 420
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
DB 421 CTGATAGACACTTATCATCAGATCCCAAGAAAGGGAATTTTATTAATGCAATTCGA 480
QY 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180

Db 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGCCCTTCGCTGGATAGCAGATAGAAAA 540
Qy 181 SerThrPheGluArgValValAlaPheAlaValGluGlyValPhePheSerGly 200
Db 541 TCTACTTTTGGGGAAGAGTGGTGGCTTTGTCTGTAGAGGAGTTTCTTCTCAGGA 600
Qy 201 SerPheAlaAlaPheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 601 TCTTTTGTCTATATTTCTGGCTAAGAAAGAGGCTTATGCCAGGACTCACTTTTTC 660
Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGGCTGATTTCCAA 720
Qy 241 TyrLeuValAsnLysProSerGluArgValArgGluIleValAspAlaValLys 260
Db 721 TACTTTAGTAAATTAAGCTTTCAAGAAAGGGTCAGGAGATCATTTGATGCTGCAAA 780
Qy 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 781 ATTGACGAGGAGTTTAAACAGAGCTTGGCAGTTGGCTCATTTGGATGAATTCATT 840
Qy 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 841 TTGATGAACACGTACATTCAGTTTGTAGCTGACAGATTACTTTGTGAACCTTGGATTC 900
Qy 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 901 AAGGTTTTTTCAGGCGAGAAATCCTTTGATTTTATGGAAGACATTTCTTTAGAGGAAA 960
Qy 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 961 ACAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTGGCAGTTATGGCAGAAACC 1020

RESULT 13

US-11-088-686-22

; Sequence 22, Application US/11088686

; GENERAL INFORMATION:

; APPLICANT: Yen, Yun

; TITLE OF INVENTION: DRUG SCREENING

; FILE REFERENCE: 14037-004001

; CURRENT APPLICATION NUMBER: US/11/088,686

; CURRENT FILING DATE: 2005-03-23

; PRIOR APPLICATION NUMBER: US 60/556,836

; PRIOR FILING DATE: 2004-03-25

; NUMBER OF SEQ ID NOS: 59

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 22

; LENGTH: 1056

; TYPE: DNA

; ORGANISM: Homo sapiens

US-11-088-686-22

Alignment Scores:

Pred. No.:	2,11e-181	Length:	1056
Score:	1813.00	Matches:	349
Percent Similarity:	100.00%	Conservative:	2
Best Local Similarity:	99.43%	Mismatches:	0
Query Match:	99.56%	Indels:	0
DB:	27	Gaps:	0

US-10-698-228-1 (1-351) x US-11-088-686-22 (1-1056)

Qy 1 MetGlyAspProGluArgProGluAlaAlaGlyLeuAspGlnAspGluArgSerSer 20
Db 1 ATGGGCGACCCGAAAGCGGAGCGGCGCGCTGGATCAGGATCAGAGATCATCTTCA 60
Qy 21 AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerArg 40

Db 61 GACACCAACGAAAGTGAATAAAGTCAATGAAGAGCCACTCTTAAGAAAGAGTTCTCGC 120
Qy 41 ArgPheValIlePheProIleGlnTyrProAspIleTrpLysMetTyrLysGlnAlaGln 60
Db 121 CGGTTTGTCTATCTTCCAATCCAGTACCCCTGATATTTGGAAAAATGTATAAACACAGC 180
Qy 61 AlaserPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys 80
Db 181 GCTTCTCTTCTGGACAGCAGAAAGAGTGCATTTATCAAGGATCTCCCTCATCTGGNACAG 240
Qy 81 LeuLeuAlaAspGluLysTyrPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp 100
Db 241 CTTTAAAGCAGATCAGAAGTACTTCTCATCTCTCACATCTTAGCCTTTTTCAGCCAGT 300
Qy 101 GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla 120
Db 301 GGAATTTGTAATGAAATTTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGTTCCAGAGCT 360
Qy 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 361 CGCTGTTTCTTGGCTTTCAAATTTCTCATCGAATGTTCTCATCAGAGATGTTAGTTTG 420
Qy 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 421 CTGATAGACACTTACATCAGATCCCAAGAAAAAGGGAATTTTATTTAATGCAATTGAA 480
Qy 161 ThrMetProTyrValLysLysLysAlaAspTrpAlaLeuArgTrpIleAlaAspArgLys 180
Db 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTGGCCCTTGCATGGATGATAGATAGAAAA 540
Qy 181 SerThrPheGlyGluArgValValAlaPheAlaAlaValGluGlyValPhePheSerGly 200
Db 541 TCTACTTTTGGGAAAGAGTGGTGGCTTTGCTGCTGAGAGGAGTTTCTTCTCAGGA 600
Qy 201 SerPheAlaAlaIlePheTrpLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 601 TCTTTTGTCTATATTTCTGGCTAAGAAAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
Qy 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 661 AATGAACCTCATCAGCAGAGATGAAGACTTCACTGTGACTTTGCTTGGCTGATTTCCAA 720
Qy 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAspAlaValLys 260
Db 721 TACTTAGTAAATAAGCCTTCAGAAAGAAAGGTCAGGAGATCATTTGATGCTGTCAAA 780
Qy 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 781 ATTGACGAGGAGTTTAAACAGAGCTTGGCAGTTGGCTCATTTGATGCTGTCAAA 840
Qy 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 841 TTGATGAACACGTACATTCAGTTTGTAGCTGACAGATTACTTTGTGAACCTTGGATTC 900
Qy 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 901 AAGGTTTTTTCAGGCGAGAAATCCTTTGATTTTATGGAAGACATTTCTTTAGAGGAAA 960
Qy 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 961 ACAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTGGCAGTTATGGCAGAAACC 1020
Qy 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1021 ACAGATAACGCTTTCACCTTGGATGCAGATTTT 1053

RESULT 14

PCT-US05-09639-24

; Sequence 24, Application PC/TUS0509639

; GENERAL INFORMATION:

; APPLICANT: Yen, Yun

; TITLE OF INVENTION: DRUG SCREENING

Qy	1	MetGlyAspProGluuArgProGluuAlaalaGlyLeuAspGlnAspGluuArgSerSer	20
Db	1	ATGGCGACCCGAAAGCCGGAAGCGCGCGGTGGATCAGGATCAGAGATCATCTTCA	60
Qy	21	AspThrAsnGluSerGluIleLysSerAsnGluGluProLeuLeuArgLysSerSerArg	40
Db	61	GACACCAACGAAAGTGAATAAAGTCAATGAAGCGCCACTCCTTAAGAAAGAGTTCTCGC	120
Qy	41	ArgPheValIlePheProIleGlnTyProAspIleTrpLysMetTyLysGlnAlaGln	60
Db	121	CGGTTTGTCTATCTTTCCAAATCCAGTACCTTGATATTGGAAAAATGATAAACAGGCACAG	180
Qy	61	AlaSerPheTrpThrAlaGluGluValAspLeuSerLysAspLeuProHisTrpAsnLys	80
Db	181	GCTTCTCTTCGACAGCAGAGAGCGTGCACCTTATCAAGAGGATCTCCCTCATCTGGAAACAAG	240
Qy	81	LeuLysAlaAspGluLysTyPheIleSerHisIleLeuAlaPhePheAlaAlaSerAsp	100
Db	241	CTTAACGACATGAAGATGCTTCTCTCACAATCTTAGCCCTTTTTCGACGCCAGTGTAT	300
Qy	101	GlyIleValAsnGluAsnLeuValGluArgPheSerGlnGluValGlnValProGluAla	120

```
Db 301 GGAATTGTAATCAAAATTTGGTGGAGCGCTTTAGTCAGGAGGTGCGAGGTTCCAGAGGCT 360
QY 121 ArgCysPheTyrGlyPheGlnIleLeuIleGluAsnValHisSerGluMetTyrSerLeu 140
Db 361 CGCTGTTTCGGGCTTTCAATTCATCGAGAAITTCACCTCAGAGATGTTAGTTG 420
QY 141 LeuIleAspThrTyrIleArgAspProLysLysArgGluPheLeuPheAsnAlaIleGlu 160
Db 421 CTGATAGACACTTACATCAGAGATCCCAAGAAAGGGAATTTTATTAATGCAATTGAA 480
QY 161 ThrMetProTyrValLysLysLysAlaAspTrrAlaLeuArgTrrIleAlaAspArgLys 180
Db 481 ACCATGCCCTATGTTAAGAAAAAGCAGATTTGGCCCTTGCATGGATAGCAGATAGAAAA 540
QY 181 SerThrPheGlyGluArgValAlaPheAlaValGluGlyValPhePheSerGly 200
Db 541 TCTACTTTTGGGAAGAGTGGTGGCTTTCCTGTAGAGGAGTTTTCITCTCAGGA 600
QY 201 SerPheAlaAlaIlePheTrrLeuLysLysArgGlyLeuMetProGlyLeuThrPheSer 220
Db 601 TCTTTTGTCTGTATATTCTGGCTAAAGAGAGAGAGTCTTATGCCAGGACTCACTTTTTC 660
QY 221 AsnGluLeuIleSerArgAspGluGlyLeuHisCysAspPheAlaCysLeuMetPheGln 240
Db 661 AATGAACCTCATCAGCAGAGATGAAGGACTTCACTGTGACTTTGCTTGCCTGATGTTCCAA 720
QY 241 TyrLeuValAsnLysProSerGluGluArgValArgGluIleIleValAlaValLys 260
Db 721 TACTTAGTAAATAAGCCTTCAGAAAGAAAGGTCAGGAGATCAATTGTGATGCTGTCAAA 780
QY 261 IleGluGlnGluPheLeuThrGluAlaLeuProValGlyLeuIleGlyMetAsnCysIle 280
Db 781 ATTGACGAGAGTTTTTAAACAGAGGCTTGGCAGTTGGCCTCATTTGGAATGAAATTGCATT 840
QY 281 LeuMetLysGlnTyrIleGluPheValAlaAspArgLeuLeuValGluLeuGlyPheSer 300
Db 841 TTGATGAAACAGTACATTCAGTTTGTAGCTGACAGATTACTTGTGAACTTGGATTCTCA 900
QY 301 LysValPheGlnAlaGluAsnProPheAspPheMetGluAsnIleSerLeuGluGlyLys 320
Db 901 AAGGTTTTTCAGGCAGAAAATCCTTTTGAATTTTATGGAACACATTTCTTTAGAAAGGAAA 960
QY 321 ThrAsnPhePheGluLysArgValSerGluTyrGlnArgPheAlaValMetAlaGluThr 340
Db 961 ACAATTTCTTTGAGAAACAGTTTCAGAGTATCAGCGTTTTGCAAGTTATGGCAGAAACC 1020
QY 341 ThrAspAsnValPheThrLeuAspAlaAspPhe 351
Db 1021 ACAGATAACGTCTTTCACCTTGGATGCAGATTTT 1053
```

Search completed: October 30, 2005, 10:06:54
Job time : 1278 secs